

23 Rec'd PCT/PTO
 31 JUL 1998
 09/117442

FORM PTO-1390 (REV 5-93)		U.S. DEPARTMENT OF COMMERCE PATENT AND TRADEMARK OFFICE		ATTORNEY DOCKET NO. P564-8013	
TRANSMITTAL LETTER TO THE UNITED STATES DESIGNATED/ELECTED OFFICE (DO/EO/US) CONCERNING A FILING UNDER 35 U.S.C. 371				DATE: 09/11/98	
				U.S. APPLN. NO. (IF KNOWN, SEE 37 CFR 1.5)	
INTERNATIONAL APPLICATION NO. PCT/EP97/00432		INTERNATIONAL FILING DATE 31 JANUARY 1997		PRIORITY DATE CLAIMED 1 FEBRUARY 1996	
TITLE OF INVENTION: RECOMBINANT EXPRESSION OF S-LAYER PROTEINS					
APPLICANT(S) FOR DO/EO/US: Werner LUBITZ, Uwe SLEYTR, Beatrix KUEN, Michaela TRUPPE, Stefan HOWORKA, Stepanka RESCH, Gerhard SCHROLL, Margit SARA					
1. <input checked="" type="checkbox"/> This is a FIRST submission of items concerning a filing under 35 U.S.C. 371. (THE BASIC FILING FEE IS ATTACHED)					
2. <input type="checkbox"/> This is a SECOND or SUBSEQUENT submission of items concerning a filing under 35 U.S.C. 371.					
3. <input checked="" type="checkbox"/> This express request to begin national examination procedures (35 U.S.C. 371(f) at any time rather than delay examination until the expiration of the applicable time limit set in 35 U.S.C. 371(b) and PCT articles 22 and 39(1).					
4. <input checked="" type="checkbox"/> A proper demand for International Preliminary Examination was made by the 19th month from the earliest claimed priority date.					
5. <input checked="" type="checkbox"/> A copy of the International Application as filed (35 U.S.C. 371(c)(2)) <ul style="list-style-type: none"> a. <input checked="" type="checkbox"/> is transmitted herewith (required only if not transmitted by the International Bureau). b. <input type="checkbox"/> has been transmitted by the International Bureau. c. <input type="checkbox"/> is not required, as the application was filed in the United States Receiving Office (RO/US) 					
6. <input checked="" type="checkbox"/> A translation of the International Application into English (35 U.S.C. 371(c)(2)).					
7. <input type="checkbox"/> Amendments to the claims of the International Application under PCT Article 19 (35 U.S.C. 371(c)(3)) <ul style="list-style-type: none"> a. <input type="checkbox"/> are transmitted herewith (required only if not transmitted by the International Bureau). b. <input type="checkbox"/> have been transmitted by the International Bureau. c. <input type="checkbox"/> have not been made; however, the time limit for making such amendments has NOT expired. d. <input type="checkbox"/> have not been made and will not be made. 					
8. <input type="checkbox"/> A translation of the amendments to the claims under PCT Article 19 (35 U.S.C. 371(c)(3)).					
9. <input checked="" type="checkbox"/> An oath or declaration of the inventor Uwe Sleyter (35 U.S.C. 371(c)(4)).					
10. <input checked="" type="checkbox"/> A translation of the annexes to the International Preliminary Examination Report under PCT Article 36 (35 U.S.C. 371(c)(5)).					
Items 11. to 16. below concern other document(s) or information included:					
11. <input checked="" type="checkbox"/> An Information Disclosure Statement under 37 CFR 1.97 and 1.98.					
12. <input type="checkbox"/> An assignment document for recording. A separate cover sheet in compliance with 37 CFR 3.28 and 3.31 is included.					
13. <input type="checkbox"/> A FIRST preliminary amendment. <input type="checkbox"/> A SECOND or SUBSEQUENT preliminary amendment.					
14. <input type="checkbox"/> A substitute specification.					
15. <input type="checkbox"/> A change of power of attorney and/or address letter.					
16. <input checked="" type="checkbox"/> Other items or information: Small Entity Declaration, PCT/IP/338, PCT/IPEA/409, PCT/IPEA/416, PCT/ISA/210 CHECK NO. 7159 Drawings - 3 sheets					

U.S. APPLN. NO. (IF KNOWN, SEE 37 C.F.R. 1.50)		INTERNATIONAL APPLICATION NO. PCT/EP97/00432		ATTORNEY DOCKET NO. P564-8013 DATE: July 31, 1998	
------------------------------------------------	--	----------------------------------------------	--	------------------------------------------------------	--

17. <input checked="" type="checkbox"/> The following fees are submitted: Basic National Fee (37 CFR 1.492(a)(1)-(5)): Search Report has been prepared by the EPO or JPO.....\$930.00 International preliminary examination fee paid to USPTO (37 CFR 1.482).....\$720.00 No international preliminary examination fee paid to USPTO (37 CFR 1.482) but international search fee paid to USPTO (37 CFR 1.445(a)(2)).....\$790.00 Neither international preliminary examination fee (37 CFR 1.482) or international search fee (37 CFR 1.445(a)(2)) paid to USPTO.....\$1,070.00 International preliminary examination fee paid to USPTO (37 CFR 1.482) and all claims satisfied provisions of PCT Article 33(2)-(4)\$ 98.00				<table border="1" style="width: 100%; border-collapse: collapse;"> <tr> <th style="width: 60%;">CALCULATIONS</th> <th style="width: 40%;">PTO USE ONLY</th> </tr> <tr><td> </td><td> </td></tr> <tr><td> </td><td> </td></tr> <tr><td> </td><td> </td></tr> <tr><td> </td><td> </td></tr> </table>		CALCULATIONS	PTO USE ONLY								
CALCULATIONS	PTO USE ONLY														

ENTER APPROPRIATE BASIC FEE AMOUNT =				\$930	
Surcharge of \$130.00 for furnishing the oath or declaration later than _ 20 _ 30 months from the earliest claimed priority date (37 CFR 1.492(e)).				\$00	
Claims	Number Filed	Number Extra	Rate		
Total Claims	- 20 =		X \$ 22.00	\$00	
Independent Claims	- 3 =		X \$ 82.00	\$00	
Multiple dependent claim(s) (if applicable)			+ \$270.00	\$00	
TOTAL OF ABOVE CALCULATIONS =				\$930	
Reduction by 1/2 for filing by small entity, if applicable. Verified Small Entity statement must also be filed. (Note 37 CFR 1.9, 1.27, 1.28).				\$465	
SUBTOTAL =				\$465	
Processing fee of \$130.00 for furnishing the English translation later the _ 20 _ 30 months from the earliest claimed priority date (37 CFR 1.492(f)).				\$00	
TOTAL NATIONAL FEE =				\$465	
Fee for recording the enclosed assignment (37 CFR 1.21(h)). The assignment must be accompanied by an appropriate cover sheet (37 CFR 3.28, 3.31). \$40.00 per property				\$00	
TOTAL FEES ENCLOSED =				\$465	
				Amount to be refunded	\$
				Charged	\$

a. ☒ A check in the amount of \$465 to cover the above fees is enclosed.
 b. ☐ Please charge my Deposit Account No. 14-1060 in the amount of \$_____ to cover the above fees. A duplicate copy of this sheet is enclosed.
 c. ☒ The Commissioner is hereby authorized to charge any additional fees which may be required, or credit any overpayment to Deposit Account No. 14-1060.

NOTE: Where an appropriate time limit under 37 CFR 1.494 or 1.495 has not been met, a petition to revive (37 CFR 1.137(a) or (b)) must be filed and granted to restore the application to pending status.

SEND ALL CORRESPONDENCE TO:
 NIKAIIDO, MARMELESTEIN, MURRAY AND ORAM
 Metropolitan Square
 655 15th Street, N.W.
 Suite 330 - G Street Lobby
 Washington, D.C. 20005-5701
 Telephone No. (202) 638-5000

Robert B. Murray
 Reg. No. 22,980

Applicant or Patentee: _____ Attorney's
Serial or Patent No.: _____ Docket No.: _____
Filed or Issued: _____
For: Recombinant expression of S-layer proteins

VERIFIED STATEMENT (DECLARATION) CLAIMING SMALL ENTITY
STATUS (37 CFR 1.9(f) and 1.27(b)) - INDEPENDENT INVENTOR

As a below named inventor, I hereby declare that I qualify as an independent inventor as defined in 37 CFR 1.9(c) for purposes of paying reduced fees under section 41(a) and (b) of Title 35, United States Code, to the Patent and Trademark Office with regard to the invention entitled Recombinant expression of S-layer proteins described in

☒ the specification filed herewith
☐ application serial no: _____, filed _____
☐ patent no. _____, issued _____

I have not assigned, granted, conveyed or licensed and am under no obligation under contract or law to assign, grant, convey or license, any rights in the invention to any person who could not be classified as an independent inventor under 37 CFR 1.9(c) if that person had made the invention, or to any concern which would not qualify as a small business concern under 37 CFR 1.9(d) or a nonprofit organization under 37 CFR 1.9(e).

Each person, concern or organization to which I have assigned, granted, conveyed, or licensed or am under an obligation under contract or law to assign, grant, convey, or license any rights in the invention is listed below:

☒ no such person, concern, or organization
☐ persons, concerns or organizations listed below*

*NOTE: Separate verified statements are required from each named person, concern or organization having rights to the invention averring to their status as small entities. (37 CFR 1.27)


FULL NAME _____
ADDRESS _____
☐ INDIVIDUAL ☐ SMALL BUSINESS CONCERN ☐ NONPROFIT ORGANIZATION

FULL NAME _____
ADDRESS _____
☐ INDIVIDUAL ☐ SMALL BUSINESS CONCERN ☐ NONPROFIT ORGANIZATION

FULL NAME _____
ADDRESS _____
☐ INDIVIDUAL ☐ SMALL BUSINESS CONCERN ☐ NONPROFIT ORGANIZATION

I acknowledge the duty to file, in this application or patent, notification of any change in status resulting in loss of entitlement to small entity status prior to paying, or at the time of paying, the earliest of the issue fee or any maintenance fee due after the date on which status as a small entity is no longer appropriate. (37 CFR 1.28(b))

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under section 1001 of Title 18 of the United States Code, and that such willful false statements may jeopardize the validity of the application, any patent issuing thereon, or any patent to which this verified statement is directed.

Uwe Sleytr
NAME OF INVENTOR _____ NAME OF INVENTOR _____ NAME OF INVENTOR _____
 _____
Signature of Inventor _____ Signature of Inventor _____ Signature of Inventor _____

09/117447

28 Rec'd PCT/PTO 02 DEC1998

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re application of:

Werner LUBITZ et al

Serial No.: 09/117,447

Filed: July 31, 1998

For: RECOMBINANT EXPRESSION OF S-LAYER PROTEINS

PRELIMINARY AMENDMENT

Assistant Commissioner
of Patents

Washington, D.C. 20231

December 2, 1998

Sir:

Prior to calculation of the filing fee and prior to the examination of this application,
please amend the above-identified application as follows:

IN THE SPECIFICATION:

Page 4, line 1, delete "pop2125" and insert therefor --pop2135--.

IN THE CLAIMS:

Please amend the claims as follows:

Claim 3, line 1, delete "or 2".

Claim 4, line 1, delete "one of the claims 1 to 3" and insert therefor --claim 1--.

Claim 13, line 1, delete "one of the claims 1 to 12" and insert therefor --claim 1--.

Claim 17, line 4, delete "or 16".

Claim 20, line 1, delete "or 19".

Claim 21, line 4, delete "or 16".

Claim 24, line 1, delete "or 23".

09/117447

Claim 30, line 4, delete "or 29".

Claims 40 and 42, line 1 of each, delete "one of the claims 37-39" and insert therefor --claim 37--.

Claim 43, line 1, delete "one of the claims 37-42" and insert therefor --claim 37--.

Please cancel claims 18, 25, 27, 31, 35 and 36 without prejudice and insert the following new claims:

- 46. Cell wherein it is transformed with a nucleic acid as claimed in claim 15.--
- 47. Cell wherein it is transformed with a vector as claimed in claim 17.--
- 48. Use of an S-layer protein as claimed in claim 21 as aa vaccine or adjuvant.--
- 49. Use of an S-layer structure as claimed in claim 22 as a vaccine or adjuvant.--
- 50. Use of an S-layer protein as claimed in claim 21 as an enzyme reactor.--
- 51. Use of an S-layer structure as claimed in claim 22 as an enzyme reactor--.
- 52. Cell wherein it is transformed with a nucleic acid as claimed in claim 28--.
- 53. Cell wherein it is transformed with a vector as claimed in claim 30.--
- 54. Use of an S-layer protein as claimed in claim 33 as a vaccine or adjuvant.--
- 55. Use of an S-layer structure as claimed in claim 34 as a vaccine or adjuvant.--
- 56. Use of an S-layer protein as claimed in claim 33 as an enzyme reactor--.
- 57. Use of an S-layer structure as claimed in claim 34 as an enzyme reactor--.

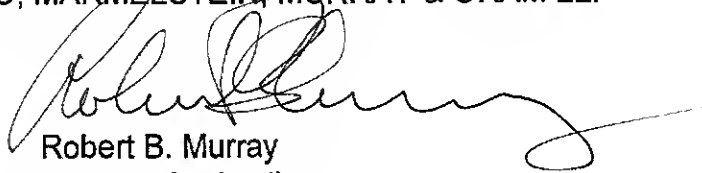
REMARKS

The above amendment corrects a typographical error in the specification. The correct E. Coli strain is identified at page 21, penultimate line. Thus it is clear that no question of new matter will arise, and entry of this amendment is in order and is requested.

The claims has been made to correct the multiple dependency of the claims and to put the application in better condition for examination.

In the event that any fees are due in connection with this paper, please charge our Deposit Account No. 14-1060.

Respectfully submitted,
NIKAIDO, MARMELSTEIN, MURRAY & ORAM LLP



Robert B. Murray
Attorney for Applicants
Reg. No. 22,980

Atty. Docket No.: P564-8013

Metropolitan Square
655 15th Street, N. W.
Suite 330 - G Street Lobby
Washington, D. C. 20005-5701
Tel (202) 638-5000
Fax (202) 638-4810

RBM/cb

3PMS

74R

31 JUL 1999

09/117447

- 1 -

Recombinant expression of S-layer proteins

Description

The present invention concerns processes for the recombinant production of S-layer proteins and modified S-layer proteins in gram-negative host cells.

Crystalline bacterial cell surface layers (S-layers) form the outermost cell wall component in many eubacteria and most of the archaebacteria (Sleytr et al. (1988), Crystalline Bacterial Cell Surface Layers, "Springer Verlag Berlin"; Messner and Sleytr, Adv. Microb. Physiol. 33 (1992), 213-275). Most of the presently known S-layer proteins are composed of identical proteins or glycoproteins which have apparent molecular weights in the range of 40,000 to 220,000. The components of S-layers are self-assembling and most of the lattices have an oblique (p2), quadratic (p4) or hexagonal (p6) symmetry. The functions of bacterial S-layers are still not completely understood but due to their location on the cell surface the porous crystalline S-layers probably serve mainly as protective coatings, molecular sieves or to promote cell adhesion and surface recognition.

Genetic data and sequence information are known for various S-layer genes from microorganisms. A review may be found in Peyret et al., Mol. Microbiol. 9 (1993), 97-109. Explicit reference is made to these data. The sequence of the sbsA gene coding for the S-layer protein of B.stearothermophilus PV72 and a process for cloning it are stated in Kuen et al. (Gene 145 (1994), 115-120).

09/117447

B.stearothermophilus PV72 is a gram-positive bacterium which is covered with a hexagonally arranged S-layer. The main component of the S-layer is a 128 kd protein which is the most frequent protein in the cell with a proportion of about 15 % relative to the total protein components. Various strains of B.stearothermophilus have been characterized which differ with regard to the type of the S-layer lattice, the molecular weight and glycosilation of the S-layer components (Messner and Sleytr (1992), supra).

The German Patent Application P 44 25 527.6 discloses the signal peptide-coding section of the S-layer gene from B.stearothermophilus and the amino acid sequence derived therefrom. The cleavage site between the signal peptide and the mature protein is located between position 30 and 31 of the amino acid sequence. The signal peptide-coding nucleic acid can be operatively linked to a protein-coding nucleic acid and can be used for the recombinant production of proteins in a process in which a transformed host cell is provided, the host cell is cultured under conditions which lead to an expression of the nucleic acid and to production and secretion of the polypeptide coded thereby and the resulting polypeptide is isolated from the culture medium. Prokaryotic organisms are preferably used as host cells in particular gram-positive organisms of the genus bacillus.

Surprisingly it was found that the recombinant production of S-layer proteins is not only possible in gram-positive prokaryotic host cells but also in gram-negative prokaryotic host cells. In this case the S-layer protein is not formed in the interior of the host cell in the form of ordered inclusion bodies but rather

unexpectedly in the form of ordered monomolecular layers.

Hence one subject matter of the present invention is a process for the recombinant production of S-layer proteins characterized in that (a) a gram-negative prokaryotic host cell is provided which is transformed with a nucleic acid coding for an S-layer protein selected from (i) a nucleic acid which comprises the nucleotide sequence shown in SEQ ID NO. 1 from position 1 to 3684 optionally without the section coding for the signal peptide, (ii) a nucleic acid which comprises a nucleotide sequence corresponding to the nucleic acid from (i) within the scope of the degeneracy of the genetic code and (iii) a nucleic acid which comprises a nucleotide sequence which hybridizes with the nucleic acids from (i) or/and (ii) under stringent conditions; (b) the host cell is cultured under conditions which lead to an expression of the nucleic acid and to production of the polypeptide coded thereby and (c) the resulting polypeptide is isolated from the host cell.

The term "stringent hybridization" is understood within the sense of the present invention to mean that a hybridization still also occurs after washing at 55°C, preferably 60°C in an aqueous low salt buffer (e.g. 0.2 x SSC) (see also Sambrook et al. (1989), Molecular Cloning. A Laboratory Manual).

The process according to the invention is carried out in gram-negative prokaryotic host cells. In this process an ordered S-layer protein structure is surprisingly obtained in the cell interior. Enterobacteria, in particular E. coli, are preferably used as host cells.

The E. coli strain pop2125 which was deposited on the 31.01.1996 at the "Deutsche Sammlung von Mikroorganismen und Zellkulturen GmbH", Mascheroder Weg 1b, D 38124 Braunschweig under the file number DSM 10509 is particularly preferred.

The process according to the invention can also be used to isolate recombinant S-layer proteins. For this one uses a nucleic acid coding for the S-layer protein which contains one or several insertions which code for peptide or polypeptide sequences. These insertions can, on the one hand, only code for peptides with a few amino acids e.g. 1-25 amino acids. On the other hand, the insertions can also code for larger polypeptides of for example up to 1000 amino acids and preferably up to 500 amino acids without loss of the ability of the S-layer protein to form a correctly folded structure. In addition to the insertions the recombinant S-layer protein can also have amino acid substitutions, in particular substitutions of individual amino acids in the region of the insertion sites as well as optionally deletions of individual amino acids or short amino acid sections of up to 30 amino acids.

Regions between the positions 1-1200 and 2200-3000 of the nucleotide sequence shown in SEQ ID NO.1 are preferred as insertion sites for polypeptide-coding sequences. Particularly preferred insertion sites are the NruI cleavage site at position 582, the PvuII cleavage site at position 878, the SnaB-I cleavage site at position 917, the PvuII cleavage site at position 2504 and the PvuII cleavage site at position 2649. It was already possible to demonstrate the insertion of a nucleic acid coding for streptavidin into the NruI cleavage site at position 581.

The peptide or polypeptide-coding insertions are preferably selected from nucleotide sequences which code for cysteine residues, regions with several charged amino acids, e.g. Arg, Lys, Asp or Glu, or Tyr residues, DNA-binding epitopes, antigenic, allergenic or immunogenic epitopes, metal-binding epitopes, streptavidin, enzymes, cytokines or antibody-binding proteins.

A particularly preferred example of an insertion into the nucleic acid coding for the S-layer protein is a nucleotide sequence coding for streptavidin. In this manner it is possible to obtain universal carrier molecules which are suitable for coupling biotinylated reagents and for detection in immunological or hybridization test procedures.

A further preferred example of insertions are antigenic, allergenic or immunogenic epitopes e.g. epitopes from pathogenic microorganisms such as bacteria, fungi, parasites etc. and viruses, or epitopes from plants or epitopes against endogenous substances e.g. cytokines as well as against toxins in particular endotoxins. Particularly preferred examples of immunogenic epitopes are epitopes from herpes viruses such as the herpes virus 6 or pseudorabies virus (Lomniczi et al., J. Virol. 49 (1984), 970-979), in particular epitopes from the genes gB, gC or/and gD, or foot-and-mouth disease virus (FMDV), in particular epitopes from the gene sections which code for VP1, VP2 or/and VP3. The immunogenic epitopes can be selected such that they promote an antibody-mediated immune reaction or/and the production of a cellular immune reaction e.g. by stimulation of T cells. Examples of suitable allergenic epitopes are birch pollen allergens e.g. Bet v I (Ebner

et al., J. Immunol. 150 (1993) 1047-1054). Antigenic epitopes are additionally particularly preferred which are able to bind and filter out endogenous or exogenous substances such as cytokines or toxins from serum or other body fluids. Such epitopes can include components of cytokine or toxin receptors or of antibodies against cytokines or toxins.

On the other hand the insertions can also code for enzymes. Preferred examples are enzymes for the synthesis of polyhydroxybutyric acid e.g. PHB synthase. Incorporation of PHB synthase into the S-layer can lead to the formation of a molecular spinning nozzle under suitable conditions when the substrate hydroxybutyric acid is provided. A further preferred example of an enzyme is bacterial luciferase. In this case when the enzyme substrate, an aldehyde, is supplied and O₂ is present, a molecular laser can be obtained.

Insertions are likewise preferred which code for cytokines such as interleukins, interferones or tumour necrosis factors. These molecules can for example be used in combination with immunogenic epitopes to prepare vaccines.

Finally insertions are also preferred which code for antibody binding proteins such as protein A or protein G or for DNA-binding or/and metal-binding epitopes such as the leucine zipper, zinc finger etc.

Thus for the first time a cell is provided by the present invention which contains immobilized recombinant polypeptides in a native form e.g. active enzymes in the cytoplasm. In this manner 50,000 - 200,000 e.g. ca.

100,000 recombinant molecules can be immobilized per m² recombinant S-layer. Up to 3000 m² S-layer can be obtained per kg recombinant E. coli cells.

In the method according to the invention the nucleic acid coding for the S-layer protein is preferably used in operative linkage with a nucleic acid coding for a signal peptide of gram-positive bacteria i.e. the signal peptide-coding nucleic acid is located on the 5' side of the S-layer protein-coding nucleic acid. Surprisingly it was found that the presence of such signal peptide sequences, which are not cleaved in the gram-negative host cells used in the invention, can improve the stability of the S-layer structures. The nucleic acid coding for the signal peptide particularly preferably comprises (a) the signal peptide-coding section of the nucleotide sequence shown in SEQ ID NO. 1, (b) a nucleotide sequence corresponding to the sequence from (a) within the scope of the degeneracy of the genetic code or/and (c) a nucleotide sequence which is at least 80 % and in particular at least 90 % homologous to the sequences from (a) or/and (b).

Yet a further subject matter of the present invention is a nucleic acid which codes for a recombinant S-layer protein and is selected from (i) a nucleic acid which comprises the nucleotide sequence shown in SEQ ID NO.1 from position 1 to 3684 optionally without the signal peptide-coding section (ii) a nucleic acid which comprises a nucleotide sequence corresponding to a nucleic acid from (i) within the scope of the degeneracy of the genetic code and (iii) a nucleic acid which comprises a nucleotide sequence which hybridizes under stringent conditions with the nucleic acids from (i) or/and (ii).

The coding nucleotide sequence of the S-layer gene sbsA from *B.stearothermophilus* including the signal peptide-coding section is shown in SEQ ID NO. 1. The signal peptide-coding section extends from position 1 to 90 of the nucleotide sequence shown in SEQ ID NO. 1. The section coding for the mature SbsA polypeptide extends from position 91 to 3684.

The sbsA gene of *B.stearothermophilus* codes for a protein with a total of 1228 amino acids including an N-terminal signal peptide with 30 amino acids (SEQ ID NO. 2). The cleavage site between the signal peptide and the mature protein is located between position 30 and 31 of the amino acid sequence. The signal peptide has a basic amino-terminal domain followed by a hydrophobic domain.

Sequence comparisons with other signal peptides indicate a certain homology to signal peptides of extracellular proteins in bacilli such as alkaline phosphatase and neutral phosphatase of *B.amyloliquefaciens* (Vasantha et al., J. Bacteriol. 159 (1984), 811-819) as well as with the signal peptides for the *B.sphaericus* gene 125 (Bowditch et al., J. Bacteriol. 171 (1989), 4178-4188) and the OWP gene of *B.brevis* (Tsuboi et al., J. Bacteriol. 168 (1986), 365-373).

A further subject matter of the present invention is a recombinant vector which contains at least one copy of a nucleic acid according to the invention. The vector is preferably replicatable in prokaryotes. The vector is particularly preferably a prokaryotic plasmid.

Yet a further subject matter of the present invention is

a host cell which is transformed with a nucleic acid or a recombinant vector according to the present invention. The cell is preferably a gram-negative prokaryotic organism and most preferably an E. coli cell. The cell according to the invention can contain a recombinant S-layer structure in its interior. Methods for the transformation of cells with nucleic acids are general state of the art (cf. Sambrook et al., supra) and therefore do not need to be elucidated.

Yet a further subject matter of the present invention is a recombinant S-layer protein which contains at least one peptide insertion or/and polypeptide insertion within the amino acid sequence shown in SEQ ID NO. 2. Preferred examples of peptide insertions and polypeptide insertions have already been elucidated.

A recombinant S-layer structure can be assembled from recombinant S-layer protein molecules according to the invention which contain at least one recombinant S-layer protein according to the invention as a subunit. Furthermore it is preferred that the S-layer structure according to the invention also contains non-modified S-layer proteins as diluent molecules. The non-modified S-layer proteins are preferably present in a molar proportion of 10-99 % relative to the total S-layer proteins.

The S-layer structure according to the invention can comprise several layers that are covalently linked together or by means of affinity binding. Covalent linkages can for example be introduced by insertions of cysteine residues and a subsequent formation of cystine bridges. Linkages by affinity binding comprise for

example antibody-antigen, antibody-protein A or antibody-protein G or streptavidin-biotin interactions.

S-layer structures which contain recombinant S-layer proteins can optionally also be prepared in a carrier-bound form. For this the S-layer structure can be reassembled from individual units in the presence of a peptidoglycan carrier to for example produce peptidoglycan layers which are covered on one or on both sides with an S-layer structure. Another method of preparing carrier-bound S-layer structures is to produce an S-layer layer at an interface between two media e.g. water/air and to immobilize this layer on a solid phase e.g. a filter membrane (cf. e.g. Pum and Sleytr (1994), Thin Solid Films 244, 882-886; Küpcü et al., (1995), Biochim. Biophys. Acta 1235, 263-269).

The recombinant S-layer proteins and S-layer structures according to the invention are suitable for a multitude of applications. An application as a vaccine or adjuvant is particularly preferred in which case recombinant S-layer proteins are used which contain immunogenic epitopes of pathogens and/or endogenous immunostimulatory polypeptides such as cytokines. In this application it is not absolutely necessary to purify the recombinant S-layer proteins. Instead they can for example be used in combination with a bacterial ghost which optionally contains additional immunogenic epitopes in its membrane.

The preparation of suitable "bacterial ghosts" is described for example in the International Patent application PCT/EP91/00967 to which reference is herewith made. In this application modified bacteria are

disclosed which are obtainable by transformation of a gram-negative bacterium with the gene of a lytically active membrane protein from bacteriophages, with the gene of a lytically active toxin release protein or with genes which contain partial sequences thereof which code for lytic proteins, culturing the bacterium, expression of this lysis gene and isolation of the resulting bacterial ghost from the culture medium.

A recombinant protein, which is obtainable by expression of a recombinant DNA in these gram-negative bacteria, can be bound to the membrane of these bacteria as described in the European Patent 0 516 655. This recombinant DNA comprises a first DNA sequence which codes for a hydrophobic, non-lytically active membrane-integrating protein domain which has an α -helical structure and is composed of 14-20 amino acids which can be flanked N- and C-terminally by 2-30 arbitrary amino acids in each case. A second DNA sequence is in operative linkage with this first DNA sequence which codes for a desired recombinant protein. Furthermore the gram-negative bacterium contains a third DNA sequence which is under a different control from the first and second DNA sequences and codes for a lytically active membrane protein from bacteriophages or a lytically active toxin release protein or for their lytically active components. So-called "bacterial ghosts" are obtained by expression and lysis of such recombinant gram-negative bacteria which contain an intact surface structure with immunogenic epitopes bound to the surface.

When these bacterial ghosts are combined with recombinant S-layers according to the invention vaccines and adjuvants can be produced which have particularly

advantageous properties.

A further particularly preferred application for recombinant S-layer proteins and S-layer structures is their use as an enzyme reactor. Such an enzyme reactor can for example be formed by a cell which contains a recombinant S-layer structure according to the invention in its interior. On the other hand the enzyme reactor can also be formed from isolated and in vitro reassembled S-layer structures or combinations of various S-layer structures.

It was found that the gram-positive bacterium *B.stearothermophilus* PV72 has an additional S-layer protein in addition to SbsA which is subsequently denoted as SbsB (Sara and Sleytr (1994), J. Bacteriol. 176, 7182-7189). It was possible to isolate and characterize the sbsB gene by amplification using suitable nucleic acid primers. The coding nucleotide sequence of the S-layer gene sbsB from *B.stearothermophilus* including the signal peptide-coding section which extends from position 1 to 93 of the nucleic acid sequence is shown in SEQ ID NO.5. The amino acid sequence derived therefrom is shown in SEQ ID NO.6. The sbsB gene codes for a protein with a total of 921 amino acids including an N-terminal signal peptide with 31 amino acids.

One subject matter of the present invention is hence a nucleic acid which codes for an S-layer protein and is selected from

- (i) a nucleic acid which comprises the nucleotide sequence from position 1 to 2763 shown in SEQ ID NO.5 optionally without the signal peptide-

- coding section,
- (ii) a nucleic acid which comprises a nucleotide sequence corresponding to the nucleic acid from (i) within the scope of the degeneracy of the genetic code and
 - (iii) a nucleic acid which comprises a nucleotide sequence that hybridizes with the nucleic acids from (i) or/and (ii) under stringent conditions.

As in the case of the sbsA gene, it is also possible to insert at least one nucleic acid insertion coding for a peptide or polypeptide into the sbsB gene within the region coding for the S-layer protein. With regard to preferred examples of insertions in the sbsB gene reference is made to the previous statements regarding the sbsA gene.

Yet a further subject matter of the present invention is a vector which contains at least one copy of an sbsB gene optionally containing an insertion. This vector can be replicated in eukaryotes, prokaryotes or in eukaryotes and prokaryotes. It can be a vector that can be integrated into the genome of the host cell or a vector which is present extrachromosomally. The vector according to the invention is preferably a plasmid in particular a prokaryotic plasmid.

Yet a further subject matter of the present invention is a host cell which is transformed with an sbsB gene wherein the sbsB gene optionally can contain an insertion. The host cell can be a eukaryotic as well as a prokaryotic cell. The cell is preferably a prokaryotic organism. Gram-positive organisms e.g. organisms of the genus bacillus as well as gram-negative organisms such

as enterobacteria in particular *E. coli* are preferred. Methods for transforming eukaryotic and prokaryotic cells with nucleic acids are known and therefore do not need to be elucidated in detail.

The present invention also concerns an SbsB protein i.e. an S-layer protein which is coded by a nucleic acid as defined above. Recombinant SbsB proteins are particularly preferred which contain one or several peptide or/and polypeptide insertions within the sbsB sequence. The SbsB part of a polypeptide according to the invention particularly preferably has a homology of at least 80 % and in particular of at least 90 % to the amino acid sequence shown in SEQ ID NO.6.

A recombinant S-layer structure can also be assembled from the recombinant SbsB-S-layer protein molecules analogous to the recombinant SbsA-S-layer structure. In this structure the non-modified S-layer proteins are preferably present in a molar proportion of 10-99 % relative to the total S-layer proteins.

The applications for the recombinant SbsB-S-layer proteins and S-layer structures according to the invention also correspond to the applications for SbsA mentioned above. In this connection its use as a vaccine or adjuvant or as an enzyme reactor is noteworthy.

Recombinant S-layer proteins are obtainable by a process in which

- (a) a host cell is provided which contains a nucleic acid coding for an S-layer protein which contains a peptide-coding or polypeptide-coding insertion within the region coding for the S-layer protein,

- (b) the host cell is cultured under conditions which lead to an expression of the nucleic acid and to production of the polypeptide coded by it and
- (c) the resulting polypeptide is isolated from the host cell or from the culture medium.

In a first preferred embodiment of this process a recombinant SbsA-S-layer protein is prepared i.e. the nucleic acid coding for the recombinant S-layer protein is selected from

- (i) a nucleic acid which comprises the nucleotide sequence from position 1 to 3684 shown in SEQ ID NO.1 optionally without the signal peptide-coding section,
- (ii) a nucleic acid which comprises a nucleotide sequence corresponding to the nucleic acid from (i) within the scope of the degeneracy of the genetic code and
- (iii) a nucleic acid which comprises a nucleotide sequence which hybridizes with the nucleic acids from (i) or/and (ii) under stringent conditions.

In a second preferred embodiment a recombinant SbsB-S-layer protein is prepared i.e. the nucleic acid coding for the recombinant S-layer protein is selected from

- (i) a nucleic acid which comprises the nucleotide sequence from position 1 to 2763 shown in SEQ ID NO.5 optionally without the signal peptide-coding section,
- (ii) a nucleic acid which comprises a nucleotide sequence corresponding to the nucleic acid from (i) within the scope of the degeneracy of the genetic code and
- (iii) a nucleic acid which comprises a nucleotide sequence which hybridizes with the nucleic acids

from (i) or/and (ii) under stringent conditions.

In addition to the recombinant SbsA and SbsB-S-layer proteins from *B.stearothermophilus* it is, however, also possible to prepare recombinant S-layer proteins from other organisms (cf. e.g. Peyret et al., (1993), supra).

The recombinant S-layer proteins can on the one hand be produced in a heterologous host cell i.e. in a host cell which originally contains no S-layer gene. Examples of such heterologous host cells are gram-negative prokaryotic organisms such as *E. coli*.

However, the heterologous expression of S-layer proteins can also take place in gram-positive prokaryotic organisms such as *B. subtilis*. For this integration vectors are preferably used which contain a native or/and a recombinant S-layer gene. When the native signal sequences are used the S-layer proteins are secreted into the culture supernatant.

However, it is often preferable to produce the recombinant S-layer proteins in homologous host cells i.e. host cells which originally contain a natural S-layer gene. In one embodiment of this homologous expression the recombinant S-layer gene is introduced into the host cell in such a way that the host cell is still able to express a further S-layer gene which codes for a non-modified S-layer protein. The non-modified S-layer protein is preferably capable of forming an S-layer structure that is compatible with the recombinant S-layer protein. An example of this embodiment of homologous expression is a *B.stearothermophilus* PV72 cell which contains intact natural sbsA genes or/and

sbsB genes and is transformed with a plasmid which contains a recombinant S-layer gene.

In a second embodiment the homologous expression can occur in a host cell in which the intact S-layer gene originally present has been inactivated. Consequently in this embodiment no further S-layer gene is expressed in the host cell which codes for a non-modified S-layer protein which is able to form a compatible S-layer structure with the recombinant S-layer protein. A specific example of such a host cell is a *B.stearothermophilus* PV72 cell in the genome of which a gene coding for a recombinant S-layer protein has been introduced, e.g. by homologous recombination, which replaces the original S-layer gene. A further example of such a host cell is a *B.stearothermophilus* cell in which the native S-layer gene has been inactivated e.g. by site-specific mutagenesis or/and homologous recombination and is transformed with a vector containing a recombinant S-layer gene.

Gram-positive prokaryotic organisms are usually used as host cells for the homologous expression of recombinant S-layer genes. *B.stearothermophilus* PV72 is particularly preferred as a host cell which can be cultured at a high temperature in a defined synthetic medium (Schuster et al., (1995), *Biotechnol. and Bioeng.* 48: 66-77).

The present invention is further elucidated by the following examples and figures.

SEQ ID NO.1 shows the complete nucleotide sequence of the coding section of the S-layer gene sbsA of *B.stearothermophilus*;

- SEQ ID NO.2 shows the amino acid sequence derived therefrom;
- SEQ ID NO.3 shows the nucleotide sequence of the primer T5-X;
- SEQ ID NO.4 shows the nucleotide sequence of the primer E;
- SEQ ID NO.5 shows the complete nucleotide sequence of the coding section of the S-layer gene sbsB of *B.stearothermophilus*;
- SEQ ID NO.6 shows the amino acid sequence derived therefrom;
- SEQ ID NO.7 shows the nucleotide sequence of a partial fragment of the streptavidin gene;
- SEQ ID NO.8 shows the nucleotide sequence of the primer NIS 2AG;
- SEQ ID NO.9 shows the nucleotide sequence of the primer LIS C3;
- Fig. 1 shows a schematic representation of the sbsA PCR fragment used to prepare the recombinant vector pBK4;
- Fig. 2 shows a schematic representation of peptide insertions in the amino acid sequence of the SbsA S-layer protein and
- Fig. 3 shows a schematic representation of amino acid substitutions and amino acid insertions in recombinant S-layer proteins.

EXAMPLES:

1. Bacterial strains, media and plasmids

Gram-positive bacteria of the strain *Bacillus stearothermophilus* PV72 were cultured at 58°C in SVIII medium (Bartelmus and Perschak, Z.Zuckerrind. 7 (1957), 276-281). Bacteria of the strain *E. coli* pop2135 (endA, thi,

hsdR, maltT, cI857, λ pR, malPQ) were cultured in LB medium (Sambrook et al., (1989), supra). Ampicillin was added to the medium at a final concentration of 100 μ g/ml to select for transformants. The plasmid pPLcAT10 (λ pL, bla, colE1) (Stanssens et al., Gene 36 (1985), 211-223) was used as the cloning vector.

2. Manipulation of DNA fragments

Restriction analysis of DNA, agarose gel electrophoresis and cloning of DNA fragments were carried out according to the standard methods described in Sambrook et al. (1989), supra.

Competent cells were transformed by electroporation using a Bio-Rad gene pulser (Bio-Rad Laboratories, Richmond, Calif. USA) according to the manufacturer's instructions.

Plasmid DNA was isolated by the method of Birnboim and Doly (Nucleic Acids Res. 7 (1979), 1513-1523). Chromosomal DNA was isolated according to the method described in Ausubel et al. (Current Protocols in Molecular Biology (1987), New York, John Wiley).

Restriction endonucleases and other enzymes were obtained from Boehringer Mannheim, New England Biolabs or Stratagene and used according to the manufacturer's instructions.

3. DNA sequencing

The DNA sequences of the 5' regions and the 3' regions

(including the region coding for the signal sequence) of the gene sbsA in the vector pPLCAT10 were determined by the dideoxy chain termination method of Sanger et al. The primers used for sequencing were constructed on the basis of the already published sbsA sequence (Kuen et al. Gene 145 (1994), 115-120).

4. PCR amplification of sbsA

The PCR amplification of the sbsA gene was carried out in a reaction volume of 100 μ l in which 200 μ M deoxynucleotides, 1 U Pfu-polymerase (Stratagene), 1 x Pfu-reaction buffer, 0.5 μ M of each oligonucleotide primer and 100 ng genomic DNA from B.stearothermophilus as a template were present. The amplification was carried out for 30 cycles in a thermocycler (Biomed thermocycler 60). Each cycle was composed of a denaturing step of 1.5 min at 95°C, an annealing step of 1 min at 56°C and 1 min at 50°C as well as an extension step of 2 min at 72°C.

The primer T5-X shown in the sequence protocol as SEQ ID NO.3 which flanks the 5' region of sbsA and contains an XbaI site and the primer E shown in the sequence protocol in SEQ ID NO.4 which flanks the 20 nucleotide upstream region of the transcription terminator of the sbsA sequence and contains a BamHI site were used as primers.

The products amplified by PCR were electrophoretically separated on a 0.8 % agarose gel and purified for cloning using the system from Gene Clean (BIO101 La Jolla, Calif. USA).

5. Cloning of the sbsA gene into the vector pPLcAT10

The sbsA gene obtained by PCR with a length of 3.79 kb was purified and cleaved with the restriction endonucleases XbaI and BamHI. The resulting XbaI-BamHI fragment was cloned into the corresponding restriction sites of the vector pPLcAT10 so that the sbsA gene was under transcriptional control of the pL promoter located upstream. The ATG start codon of the sbsA sequence was reconstructed by the cloning procedure. The cloned sbsA sequence contained the N-terminal signal sequence of sbsA and ended 20 nt after the transcription terminator. After ligation of the vector DNA with the sbsA fragment, the E. coli strain pop2135 was transformed by electrotransformation. The resulting clones were subjected to a DNA restriction analysis. A positive clone was sequenced in order to verify the correct sequence transitions at the 5' and 3' ends. This clone was named pBK4.

A schematic representation of the 3.79 kb XbaI sbsA fragment and its location in the multiple cloning site of the plasmid pBK4 is shown in Fig. 1 (abbreviations: tT: transcription terminator; ori: origin of the DNA replication; amp: ampicillin resistance gene).

6. Recombinant expression of the sbsA gene in E. coli

E. coli pop2135/pBK4 cells were cultured at 28°C until an optical density OD₆₀₀ of 0.3 was reached. Then the expression of sbsA was induced by increasing the culture temperature from 28°C to 42°C. 1.5 ml aliquots were taken before and 1, 2, 3 and 5 hours after induction of the sbsA expression. E. coli pop2135/pPLcAT10 (cultured under the same conditions) and B.stearothermophilus PV72

were used as controls.

Culture supernatants and cell extracts from all samples were examined for the expression of S-layer proteins by SDS-PAGE and Western immunoblotting.

An additional strong protein band with the same molecular weight as the wild type SbsA protein was found in extracts from *E. coli* cells transformed with pBK4. No degradation products of SbsA itself were found in a period of up to 5 hours after induction of expression. Thus presumably the S-layer protein sbsA is stable in *E. coli* and is not degraded by proteases.

A densitometric determination of the relative amount of SbsA protein was carried out. At a time point of 4 hours after induction the sbsA protein was in a proportion of ca. 16 % relative to the total cellular protein.

The SbsA protein produced in *E. coli* migrated in the SDS gel slightly more slowly than the natural SbsA protein from *B. stearothermophilus*. Experiments to determine the N-terminal amino acid sequence of the SbsA protein by Edman degradation were not successful due to a blocking of the N-terminus. Thus presumably the signal sequence was not cleaved in *E. coli*.

A Western blot analysis of total cell extracts and culture supernatants of *E. coli*/pBK4 also only yielded a single sbsA-specific protein band with a slightly higher molecular weight than wild type SbsA protein from *stearothermophilus*.

For the Western blot the proteins were transferred onto a nitrocellulose membrane and incubated with a polyclonal antiserum against SbsA from rabbits. The preparation of this antiserum is described in Egelseer et al. (J. Bacteriol. 177 (1995), 1444-1451). A conjugate of goat anti-rabbit IgG and alkaline phosphatase was used to detect bound SbsA-specific antibodies.

No SbsA protein could be detected from supernatants from E. coli cells transformed with pBK4 even after induction of sbsA gene expression. This shows that SbsA is not exported into the surrounding medium.

7. Location and organisation of the S-layer protein SbsA in the cytoplasm of E. coli

Cells of E. coli pop2135/pBK4 which were harvested from cultures 1, 2, 3 and 5 hours after induction of the S-layer protein expression were examined for the intracellular organisation of sbsA. Non-induced cells cultured at 28°C and cells of B.stearothermophilus PV72 were examined as controls.

For this whole cells of both organisms were fixed and embedded in detection resin according to the method of Messner et al. (Int. J.Syst.Bacteriol. 34 (1984), 202-210). Subsequently ultrathin sections of the embedded preparations were prepared and stained with uranyl acetate.

The cytoplasm of non-induced E. coli cells exhibited the typical granular structure which did not change even when the OD of the suspensions increased. Longitudinal

sections of *E. coli* cells which were harvested 1 hour after induction of the S-layer protein expression exhibited parallel, leaf-like structures in the cytoplasm. From cross sections it was apparent that these structures have a concentric arrangement.

The amount of leaf-like structures considerably increased between 1 and 2 hours after induction of the *sbsA* expression and afterwards remained essentially constant.

The *sbsA* protein recombinantly produced in *E. coli* could also be detected by immunogold labelling with *sbsA*-specific antibodies. An ordered structure of the recombinantly produced SbsA protein was also found with this detection method.

It was clearly apparent from these morphological data that the SbsA protein did not aggregate to form irregular inclusion bodies but rather formed monomolecular S-layer crystals. A remarkable property of the SbsA-S-layer layers assembled in *E. coli* was the concentric arrangement at defined distances. The presence of the signal sequence did not interfere with correct assembly.

8. Preparation of recombinant *sbsA*-S-layer genes

8.1 Insertion of a 6 bp long DNA sequence

A modified kanamycin cassette (1.3 kb) was used for the site-specific insertion mutagenesis of the *sbsA* gene which was isolated by cleavage of the plasmid pWJC3

(obtained from W.T. McAllister, New York) by SmaI. The cassette was ligated into five different blunt-ended restriction sites of the sbsA gene, i.e. into the NruI site at position bp 582 (pSL582), into the SnaBI site at position bp 917 (pSL917) and into each of the PvuII sites at positions bp 878 (pSL878), bp 2504 (pSL2504) and bp 2649 (pSL2649). After selection of kanamycin-resistant clones, the cassette was removed from the insertion site by cleavage with ApaI followed by a religation of the S-layer plasmid pBK4. The cutting out and religation procedure left an insertion of 6 bp CCCGGG (ApaI restriction site). The system of this linker insertion is shown schematically in Fig. 2.

The resulting recombinant S-layer genes code for modified sbsA proteins elongated by 2 amino acids.

The specific changes in the primary structure of the sbsA proteins are shown in Fig. 3. In the clone pSL582 the insertion led to the incorporation of glycine and proline between the amino acids 194 and 195 at the N-terminus of the SbsA protein. The amino acids alanine and arginine were inserted in the clone pSL917 between the amino acids 306 and 307. In the clone pSL2649 glycine and proline were inserted between the amino acids at positions 883 and 884. An insertion of alanine and proline between the amino acids 293 and 294 was obtained in the clone pSL878. Furthermore the alanine at position 293 was substituted by glycine. In the clone pSL2504 the amino acids alanine and proline were inserted between the amino acids 835 and 836 and the alanine at position 835 was replaced by glycine.

All clones obtained by insertion mutagenesis retained

their ability to synthesise the S-layer protein.

In order to test the ability of the modified proteins to assemble into S-layer structures, ultrathin longitudinal sections of whole cells which had been cultured for 4 hours under inductive conditions were prepared according to the procedure described in section 7. It was found that the cytoplasm of all five clones is filled with parallel, leaf-like structures which follow the curve of the cell poles. There were no morphological differences of the cytoplasm in the 5 different clones examined. Exactly the same leaf-like structures were found as in the assembly of the wild type SbsA protein in *E. coli* (section 7).

8.2 Insertion of a DNA sequence coding for streptavidin

In order to examine whether the insertion of larger protein sequences into the SbsA protein can also be tolerated, a DNA fragment coding for a part of streptavidin (160 amino acids) provided with ApaI linkers (SEQ ID NO.7) was gene inserted into the ApaI restriction site of the sbsA clones pSL582, pSL878, pSL917 and pSL2649 prepared in the example on page 1. The streptavidin sequence was inserted in SL582 in the codon 197, in pSL878 between codon 295 and 296, in pSL917 in the codon 308 and 309 and in pSL2649 in the codon 886. It was possible to detect the expression of SbsA-streptavidin fusion proteins in all constructs by SDS-PAGE and immunoblots. It was found by EM analysis that a self assembly of the S-layer structure was possible in the fusion proteins containing insertions in the codon 197 and between the codons 295 and 296.

The SbsA-streptavidin fusion proteins can be isolated as monomers and reassembled to form homogeneous SbsA-streptavidin S-layers or mixed SbsA-streptavidin/SbsA-S-layers. They can be used to bind biotinylated substances as well as to determine the binding capacity of enzymes and other bound molecules.

8.3 Insertion of a DNA sequence coding for BetvI

A DNA sequence coding for the open reading frame of BetvI (161 amino acids) the main pollen allergen of the birch (Ferreira et al., J. Biol. Chem. 268 (1993), 19574-19580) was inserted at the ApaI site into the sbsA clone pSL878. It was possible to detect the expression of an SbsA-BetvI fusion protein which contained an immunologically active BetvI domain.

The resulting fusion protein can be used for therapeutic or diagnostic purposes. Hence it can be attempted by administration of the fusion protein to convert a T_H2 -directed IgE antibody reaction into a T_H1 -mediated reaction against BetvI. In this manner it is possible to suppress the occurrence of symptoms of a pollen allergy. Furthermore SbsA-BetvI fusion proteins can be used to test for anti-BetvI antibody concentrations or/and to reduce high concentrations of anti-BetvI IgE.

8.4 Insertion of a DNA sequence coding for a pseudorabies virus antigen

The DNA sequence coding for the gB epitope SmaBB (255 amino acids) (nucleotides 489-1224 corresponding to the coordinates according to the EMBL-Seq: HEHSSGP2) from the pseudorabies virus was inserted into SSpI site of

the sbsA gene after nt 3484 (between codon 1161 and 1162). It was possible to detect the expression of SbsA-SmaBB fusion proteins.

The fusion proteins can be used to test gB-specific immune reactions. A Western blot analysis using a monoclonal antibody which corresponds to the inserted sequence showed the immunological activity of the viral domain within the recombinant SbsA-SmaBB proteins.

8.5 Insertion of a DNA sequence coding for the PHB synthase (PhbC) from *Alcaligenes eutrophus* H16

A regular arrangement of polypeptide structures with enzymatic activity on the surface of S-layers is an important goal in the production of immobilized enzymes within a living cell and in the case of the 590 amino acid long PHB synthase for the production of a molecular machine for biopolymer synthesis.

The phbC gene was isolated by PCR from the plasmid p4A (Janes et al., Molecular characterisation of the poly- β -hydroxy-butyrate biosynthesis in *Alcaligenes eutrophus* H16. In: Novel Biodegradable Microbial Polymers (publisher Daves, E.A.), pp 175-190 (1990), Kluwer, Dordrecht) as a 1770 nt long DNA fragment (corresponding to an open reading frame of 590 amino acids) and inserted into the ApaI cleavage site of the sbsA clone pSL878 to obtain the plasmid pSbsA-PhbC. It was possible to detect the expression of an SbsA-PhbC fusion protein of ca. 195 kD in an *E. coli* cell transformed with this plasmid. When two copies of the phbC gene were inserted one behind the other into the ApaI site of pSL878, it was possible to detect the expression of a fusion

protein of ca. 260 kD.

For a functional test of the enzymatic activity of the SbsA-PhbC construct, the *E. coli* cells which contained the plasmid pSbsA-PhbC were co-transformed with the plasmid pUMS which contains the β -ketothiolase (PhbA) and the acetoacetyl-CoA reductase (PhbB) from *A. eutrophus* (Kalousek et al., Genetic engineering of PHB-synthase from *Alcaligenes eutrophus* H16. In: Proceedings of the International Symposium on Bacterial Polyhydroxyalkanoates, pp 426-427 (1993), publisher Schlegel H. G., Steinbüchel A. Goltze Press, Göttingen). The poly- β -hydroxybutyrate formation in the co-transformed *E. coli* cells was detectable by staining with Sudan black, gas chromatography and electron microscopy. These findings show that the SbsA-PhbC construct is enzymatically active and represents a successful example of the immobilization of enzymes on intracellular S-layer matrices.

8.6 Insertion of a DNA sequence coding for a bacterial luciferase gene

A monocistronic LuxAB gene with a length of 2,070 nt which contains the fusion protein LuxAB composed of the two subunits LuxB and LuxA of the bacterial luciferase from *Vibrio harveyi* was isolated from the plasmid pT7-mut3 (Boylan et al., J. Biol. Chem. 264 (1989), 1915-1918) by PCR and inserted into the *Apa*I site of the clone pSL878 prepared in example 8.1 to obtain the plasmid pBK878-LuxAB. It was possible to detect the expression of an SbsA-PhbC fusion protein of ca. 207 kD in an *E. coli* cell transformed with this plasmid. The enzymatic activity of the fusion protein was

demonstrated by the method described in Boylan et al.,
Supra.

9. Isolation and characterization of the sbsB gene

The basis for the isolation of the sbsB gene was the amino acid sequence of the N-terminus as well as the sequence of three internal peptides of the SbsB protein. Starting with these peptide sequences, degenerate oligonucleotide primers were constructed and used for the PCR. In this manner a 1076 bp long PCR fragment from the chromosomal DNA of B.stearothermophilus was amplified, cloned and sequenced (corresponding to position 100-1176 of the sequence shown in SEQ ID NO.5).

The method of inverse PCR was used to amplify the sections on the 5' side and 3' side of the sbsB gene and stepwise overlapping DNA fragments were obtained with the aid of various primer combinations and sequenced.

The primer NIS 2AG shown in the sequence protocol as SEQ ID NO.8 which contains the 5' region of sbsB as well as the primer LIS C3 shown in the sequence protocol of SEQ ID NO.9 which contains the 3' region of sbsB were used as primers to amplify the complete sbsB gene.

The PCR fragment obtained in this manner which contains the nucleotide sequence shown in SEQ ID NO.5 with 5' and 3' BamHI restriction cleavage sites was cloned as described in example 5 into the vector pPLCAT10 in which the expression takes place under the control of the lambda PL promoter.

Furthermore the sbsB-PCR fragment with the 5' side EcoRI and 3' side BamHI cleavage site were cloned into the vector pUC18 in which the expression took place under the control of the lac promoter.

The detection of the sbsB expression was carried out as described in examples 6 and 7 by SDS gel electrophoresis and electron microscopy.

10. Preparation of recombinant sbsB-S-layer genes

Recombinant sbsB genes were prepared analogously to the methods described in example 8.

Thus in accordance with the method described in example 8.1, a 6 nt long DNA sequence containing an ApaI restriction cleavage site was introduced at various positions into the sbsB-layer gene. The recombinant sbsB clones pAK407, pAK481 and pAK1582 with ApaI cleavage sites at nt 407 (codon 136), 481 (codon 161/162) and 1582 (codon 528/529) were obtained in this manner. These clones obtained by insertion mutagenesis retained their ability to synthesize the S-layer protein and form S-layer structures.

Analogously to the method described in example 8.2, a DNA fragment coding for streptavidin was inserted into the ApaI restriction sites of the sbsB clones pAK407 and pAK481.

Analogously to example 8.4, a DNA sequence coding for the gB epitope SmaBB was inserted into the ApaI cleavage sites of the sbsB clones pAK481 and pAK1582. It was

possible to detect the expression of sbsB-SmaB fusion proteins of ca. 130 kD in the E. coli cells transformed with the resulting recombinant plasmids. When two copies of the SmaBB epitopes were inserted one behind the other into the ApaI cleavage site of pAK481 it was possible to detect the expression of a fusion protein of ca. 157 kD. The SmaBB domains of the fusion proteins were recognized by specific antibodies.

Analogously to example 8.6 it was possible to detect the expression of a 175 kD SbsB-LuxAB fusion protein when the LuxAB sequence was inserted into the ApaI cleavage site of pAK407.

11. Heterologous expression of sbsA and sbsB in Bacillus subtilis

The integration vector pX (Kim, L., Mogk, A. and Schumann W., Gene 181 (1996), 71-76: A xylose-inducible Bacillus subtilis integration vector and its application) was used for the heterologous expression of sbsA and sbsB in B. subtilis. The S-layer genes in the resulting recombinant expression vectors are under the transcriptional control of the xyl promoter.

Transformants of B. subtilis containing an S-layer gene integrated in the chromosome exhibited an expression of large amounts of S-layer proteins in the supernatant of the cells which was inducible by addition of xylose to the growth medium. This shows that the signal sequences of sbsA and sbsB are recognized by the B. subtilis cell.

In an analogous manner it was possible to achieve a heterologous expression of recombinant sbsA and sbsB layer genes in B. subtilis.

(1) GENERAL INFORMATION:

(A) NAME: Werner Lubitz
(B) ROAD: Schoenborngasse 12/7
(C) CITY: Wien
(E) COUNTRY: Austria
(F) ZIP CODE: 1080

(ii) TITLE OF INVENTION: Recombinant expression of
S-layer proteins

(iv) COMPUTER READABLE FORM:

(2) INFORMATION FOR SEQ ID NO: 1:

(A) LENGTH: 3687 base pairs
(B) TYPE: nucleotide
(C) STRANDEDNESS: both
(D) TOPOLOGY: linear

(A) ORGANISM: Bacillus stearothermophilus
(B) STRAIN: PV72

(vii) IMMEDIATE ORIGIN:
(B) CLONE(S): sbsA

```
(ix) CHARACTERISTIC:
      (A) NAME/KEY: CDS
      (B) POSITION: 1..3684
```

```
(ix) CHARACTERISTIC:
      (A) NAME/KEY: sig_peptide
      (B) POSITION: 1..90
```

(ix) CHARACTERISTIC:

(A) NAME/KEY: mat_peptide

(B) POSITION: 91..3684

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

10

15	ATG	GAT	AGG	AAA	AAA	GCT	GTG	AAA	CTA	GCA	ACA	GCA	AGT	GCT	ATT	GCA	48
	Met	Asp	Arg	Lys	Lys	Ala	Val	Lys	Leu	Ala	Thr	Ala	Ser	Ala	Ile	Ala	
	-30					-25				-20						-15	
	GCA	AGT	GCA	TTT	GTC	GCT	GCA	AAT	CCA	AAC	GCT	TCT	GAA	GCG	GCT	ACA	96
	Ala	Ser	Ala	Phe	Val	Ala	Ala	Asn	Pro	Asn	Ala	Ser	Glu	Ala	Ala	Thr	
				-10					-5						1		
20	GAT	GTA	GCA	ACA	GTA	GTA	AGC	CAA	GCA	AAA	GCA	CAG	TTC	AAA	AAA	GCA	144
	Asp	Val	Ala	Thr	Val	Val	Ser	Gln	Ala	Lys	Ala	Gln	Phe	Lys	Lys	Ala	
			5					10					15				
25	TAC	TAT	ACT	TAC	AGC	CAT	ACA	GTA	ACG	GAA	ACT	GGT	GAA	TTC	CCA	AAC	192
	Tyr	Tyr	Thr	Tyr	Ser	His	Thr	Val	Thr	Glu	Thr	Gly	Glu	Phe	Pro	Asn	
		20					25					30					
30	ATT	AAC	GAT	GTA	TAT	GCT	GAA	TAC	AAC	AAA	GCG	AAA	AAA	CGA	TAC	CGT	240
	Ile	Asn	Asp	Val	Tyr	Ala	Glu	Tyr	Asn	Lys	Ala	Lys	Lys	Arg	Tyr	Arg	
	35					40					45					50	
35	GAT	GCG	GTA	GCA	TTA	GTG	AAT	AAA	GCA	GGT	GGC	GCG	AAA	AAA	GAC	GCT	288
	Asp	Ala	Val	Ala	Leu	Val	Asn	Lys	Ala	Gly	Gly	Ala	Lys	Lys	Asp	Ala	
					55					60					65		
40	TAC	TTA	GCT	GAT	TTA	CAA	AAA	GAA	TAT	GAA	ACT	TAC	GTT	TTC	AAA	GCA	336
	Tyr	Leu	Ala	Asp	Leu	Gln	Lys	Glu	Tyr	Glu	Thr	Tyr	Val	Phe	Lys	Ala	
				70					75					80			
45	AAC	CCT	AAA	TCT	GGC	GAA	GCT	CGT	GTA	GCA	ACT	TAC	ATC	GAT	GCT	TAC	384
	Asn	Pro	Lys	Ser	Gly	Glu	Ala	Arg	Val	Ala	Thr	Tyr	Ile	Asp	Ala	Tyr	
			85					90					95				
50	AAC	TAT	GCA	ACA	AAA	TTA	GAC	GAA	ATG	CGC	CAA	GAG	CTA	GAG	GCT	GCT	432
	Asn	Tyr	Ala	Thr	Lys	Leu	Asp	Glu	Met	Arg	Gln	Glu	Leu	Glu	Ala	Ala	
		100					105					110					
55	GTT	CAA	GCA	AAA	GAT	TTA	GAA	AAA	GCA	GAA	CAA	TAC	TAT	CAC	AAA	ATT	480
	Val	Gln	Ala	Lys	Asp	Leu	Glu	Lys	Ala	Glu	Gln	Tyr	Tyr	His	Lys	Ile	
	115					120				125						130	
60	CCT	TAT	GAA	ATT	AAA	ACT	CGC	ACA	GTC	ATT	TTA	GAT	CGC	GTA	TAT	GGT	528
	Pro	Tyr	Glu	Ile	Lys	Thr	Arg	Thr	Val	Ile	Leu	Asp	Arg	Val	Tyr	Gly	
					135					140					145		
65	AAA	ACA	ACT	CGT	GAT	TTA	CTT	CGC	TCT	ACA	TTT	AAA	GCA	AAA	GCA	CAA	576
	Lys	Thr	Thr	Arg	Asp	Leu	Leu	Arg	Ser	Thr	Phe	Lys	Ala	Lys	Ala	Gln	
				150					155					160			
70	GAA	CTT	CGC	GAC	AGC	TTA	ATT	TAT	GAT	ATT	ACC	GTT	GCA	ATG	AAA	GCG	624
	Glu	Leu	Arg	Asp	Ser	Leu	Ile	Tyr	Asp	Ile	Thr	Val	Ala	Met	Lys	Ala	
			165					170					175				

65

	CGC	GAA	GTA	CAA	GAC	GCT	GTG	AAA	GCA	GGC	AAT	TTA	GAC	AAA	GCT	AAA	672
	Arg	Glu	Val	Gln	Asp	Ala	Val	Lys	Ala	Gly	Asn	Leu	Asp	Lys	Ala	Lys	
	180						185					190					
5	GCT	GCT	GTT	GAT	CAA	ATC	AAT	CAA	TAC	TTA	CCA	AAA	GTA	ACA	GAT	GCT	720
	Ala	Ala	Val	Asp	Gln	Ile	Asn	Gln	Tyr	Leu	Pro	Lys	Val	Thr	Asp	Ala	
	195					200					205					210	
10	TTC	AAA	ACT	GAA	CTA	ACA	GAA	GTA	GCG	AAA	AAA	GCA	TTA	GAT	GCA	GAT	768
	Phe	Lys	Thr	Glu	Leu	Thr	Glu	Val	Ala	Lys	Lys	Ala	Leu	Asp	Ala	Asp	
					215					220					225		
15	GAA	GCT	GCG	CTT	ACT	CCA	AAA	GTT	GAA	AGT	GTA	AGT	GCG	ATT	AAC	ACT	816
	Glu	Ala	Ala	Leu	Thr	Pro	Lys	Val	Glu	Ser	Val	Ser	Ala	Ile	Asn	Thr	
				230					235					240			
20	CAA	AAC	AAA	GCT	GTT	GAA	TTA	ACA	GCA	GTA	CCA	GTG	AAC	GGA	ACA	CTA	864
	Gln	Asn	Lys	Ala	Val	Glu	Leu	Thr	Ala	Val	Pro	Val	Asn	Gly	Thr	Leu	
			245					250					255				
25	AAA	TTA	CAA	CTT	TCA	GCT	GCT	GCA	AAT	GAA	GAT	ACA	GTA	AAC	GTA	AAT	912
	Lys	Leu	Gln	Leu	Ser	Ala	Ala	Asn	Glu	Asp	Thr	Val	Val	Asn	Val	Asn	
		260					265					270					
30	ACT	GTA	CGT	ATC	TAT	AAA	GTG	GAC	GGT	AAC	ATT	CCA	TTT	GCC	CTT	AAT	960
	Thr	Val	Arg	Ile	Tyr	Lys	Val	Asp	Gly	Asn	Ile	Pro	Phe	Ala	Leu	Asn	
						280					285					290	
35	ACG	GCA	GAT	GTT	TCT	TTA	TCT	ACA	GAC	GGA	AAA	ACT	ATC	ACT	GTG	GAT	1008
	Thr	Ala	Asp	Val	Ser	Leu	Ser	Thr	Asp	Gly	Lys	Thr	Ile	Thr	Val	Asp	
					295					300					305		
40	GCT	TCA	ACT	CCA	TTC	GAA	AAT	AAT	ACG	GAG	TAT	AAA	GTA	GTA	GTT	AAA	1056
	Ala	Ser	Thr	Pro	Phe	Glu	Asn	Asn	Thr	Glu	Tyr	Lys	Val	Val	Val	Lys	
				310					315					320			
45	GGT	ATT	AAA	GAC	AAA	AAT	GCG	AAA	GAA	TTT	AAA	GAA	GAT	GCA	TTC	ACT	1104
	Gly	Ile	Lys	Asp	Lys	Asn	Gly	Lys	Glu	Phe	Lys	Glu	Asp	Ala	Phe	Thr	
			325				330						335				
50	TTC	AAG	CTT	CGA	AAT	GAT	GCT	GTA	GTT	ACT	CAA	GTG	TTT	GGA	ACT	AAT	1152
	Phe	Lys	Leu	Arg	Asn	Asp	Ala	Val	Val	Thr	Gln	Val	Phe	Gly	Thr	Asn	
			340				345					350					
55	GTA	ACA	AAC	AAC	ACT	TCT	GTA	AAC	TTA	GCA	GCA	GGT	ACT	TTC	GAC	ACT	1200
	Val	Thr	Asn	Asn	Thr	Ser	Val	Asn	Leu	Ala	Ala	Gly	Thr	Phe	Asp	Thr	
						360					365					370	
60	GAC	GAT	ACT	TTA	ACA	GTA	GTA	TTT	GAT	AAG	TTG	TTA	GCA	CCT	GAA	ACT	1248
	Asp	Asp	Thr	Leu	Thr	Val	Val	Phe	Asp	Lys	Leu	Leu	Ala	Pro	Glu	Thr	
					375					380					385		
65	GTA	AAC	AGC	TCG	AAC	GTT	ACT	ATT	ACA	GAT	GTT	GAA	ACT	GGA	AAA	CGC	1296
	Val	Asn	Ser	Ser	Asn	Val	Thr	Ile	Thr	Asp	Val	Glu	Thr	Gly	Lys	Arg	
				390					395					400			
70	ATT	CCA	GTA	ATT	GCA	TCT	ACT	TCT	GGT	TCT	ACA	ATT	ACT	ATT	ACG	TTA	1344
	Ile	Pro	Val	Ile	Ala	Ser	Thr	Ser	Gly	Ser	Thr	Ile	Thr	Ile	Thr	Leu	
				405				410					415				
75	AAA	GAA	GCG	TTA	GTA	ACT	GGT	AAA	CAA	TAT	AAA	CTT	GCT	ATC	AAT	AAT	1392
	Lys	Glu	Ala	Leu	Val	Thr	Gly	Lys	Gln	Tyr	Lys	Leu	Ala	Ile	Asn	Asn	
							425					430					
80	GTT	AAA	ACA	TTA	ACT	GGT	TAC	AAT	GCA	GAA	GCT	TAC	GAG	TTA	GTG	TTC	1440
	Val	Lys	Thr	Leu	Thr	Gly	Tyr	Asn	Ala	Glu	Ala	Tyr	Glu	Leu	Val	Phe	
						440					445					450	

00117447 100000

	ACT	GCA	AAC	GCA	TCA	GCA	CCA	ACT	GTT	GCT	ACC	GCT	CCT	ACT	ACT	TTA	1488
	Thr	Ala	Asn	Ala	Ser	Ala	Pro	Thr	Val	Ala	Thr	Ala	Pro	Thr	Thr	Leu	
				455						460					465		
5	GGT	GGT	ACA	ACT	TTA	TCT	ACT	GGT	TCT	CTT	ACA	ACA	AAT	GTT	TGG	GGT	1536
	Gly	Gly	Thr	Thr	Leu	Ser	Thr	Gly	Ser	Leu	Thr	Thr	Asn	Val	Trp	Gly	
				470					475					480			
10	AAA	TTG	GCT	GGT	GGT	GTG	AAT	GAA	GCT	GGA	ACT	TAT	TAT	CCT	GGT	CTT	1584
	Lys	Leu	Ala	Gly	Gly	Val	Asn	Glu	Ala	Gly	Thr	Tyr	Tyr	Pro	Gly	Leu	
			485					490					495				
15	CAA	TTC	ACA	ACA	ACG	TTT	GCT	ACT	AAG	TTA	GAC	GAA	TCT	ACT	TTA	GCT	1632
	Gln	Phe	Thr	Thr	Thr	Phe	Ala	Thr	Lys	Leu	Asp	Glu	Ser	Thr	Leu	Ala	
		500					505					510					
20	GAT	AAC	TTT	GTA	TTA	GTT	GAA	AAA	GAA	TCT	GGT	ACA	GTT	GTT	GCT	TCT	1680
	Asp	Asn	Phe	Val	Leu	Val	Glu	Lys	Glu	Ser	Gly	Thr	Val	Val	Ala	Ser	
	515					520					525					530	
25	GAA	CTA	AAA	TAT	AAT	GCA	GAC	GCT	AAA	ATG	GTA	ACT	TTA	GTG	CCA	AAA	1728
	Glu	Leu	Lys	Tyr	Asn	Ala	Asp	Ala	Lys	Met	Val	Thr	Leu	Val	Pro	Lys	
					535					540					545		
30	GCG	GAC	CTT	AAA	GAA	AAT	ACA	ATC	TAT	CAA	ATC	AAA	ATT	AAA	AAA	GGC	1776
	Ala	Asp	Leu	Lys	Glu	Asn	Thr	Ile	Tyr	Gln	Ile	Lys	Ile	Lys	Lys	Gly	
				550					555				560				
35	TTG	AAG	TCC	GAT	AAA	GGT	ATT	GAA	TTA	GGC	ACT	GTT	AAC	GAG	AAA	ACA	1824
	Leu	Lys	Ser	Asp	Lys	Gly	Ile	Glu	Leu	Gly	Thr	Val	Asn	Glu	Lys	Thr	
			565					570					575				
40	TAT	GAG	TTC	AAA	ACT	CAA	GAC	TTA	ACT	GCT	CCT	ACA	GTT	ATT	AGC	GTA	1872
	Tyr	Glu	Phe	Lys	Thr	Gln	Asp	Leu	Thr	Ala	Pro	Thr	Val	Ile	Ser	Val	
		580					585					590					
45	ACG	TCT	AAA	AAT	GGC	GAC	GCT	GGA	TTA	AAA	GTA	ACT	GAA	GCT	CAA	GAA	1920
	Thr	Ser	Lys	Asn	Gly	Asp	Ala	Gly	Leu	Lys	Val	Thr	Glu	Ala	Gln	Glu	
	595				600						605					610	
50	TTT	ACT	GTG	AAG	TTC	TCA	GAG	AAT	TTA	AAT	ACA	TTT	AAT	GCT	ACA	ACC	1968
	Phe	Thr	Val	Lys	Phe	Ser	Glu	Asn	Leu	Asn	Thr	Phe	Asn	Ala	Thr	Thr	
					615					620					625		
55	GTT	TCG	GGT	AGC	ACA	ATC	ACA	TAC	GGT	CAA	GTT	GCT	GTA	GTA	AAA	GCG	2016
	Val	Ser	Gly	Ser	Thr	Ile	Thr	Tyr	Gly	Gln	Val	Ala	Val	Val	Lys	Ala	
				630					635					640			
60	GGT	GCA	AAC	TTA	TCT	GCT	CTT	ACA	GCA	AGT	GAC	ATC	ATT	CCA	GCT	AGT	2064
	Gly	Ala	Asn	Leu	Ser	Ala	Leu	Thr	Ala	Ser	Asp	Ile	Ile	Pro	Ala	Ser	
			645					650					655				
65	GTT	GAA	GCG	GTT	ACT	GGT	CAA	GAT	GGA	ACA	TAC	AAA	GTG	AAA	GTT	GCT	2112
	Val	Glu	Ala	Val	Thr	Gly	Gln	Asp	Gly	Thr	Tyr	Lys	Val	Lys	Val	Ala	
		660					665					670					
70	GCT	AAC	CAA	TTA	GAA	CGT	AAC	CAA	GGG	TAC	AAA	TTA	GTA	GTG	TTC	GGT	2160
	Ala	Asn	Gln	Leu	Glu	Arg	Asn	Gln	Gly	Tyr	Lys	Leu	Val	Val	Phe	Gly	
	675					680					685					690	
75	AAA	GGT	GCA	ACA	GCT	CCT	GTT	AAA	GAT	GCT	GCA	AAT	GCA	AAT	ACT	TTA	2208
	Lys	Gly	Ala	Thr	Ala	Pro	Val	Lys	Asp	Ala	Ala	Asn	Ala	Asn	Thr	Leu	
					695					700					705		
80	GCA	ACT	AAC	TAT	ATC	TAT	ACA	TTT	ACA	ACT	GAA	GGT	CAA	GAC	GTA	ACA	2256
	Ala	Thr	Asn	Tyr	Ile	Tyr	Thr	Phe	Thr	Thr	Glu	Gly	Gln	Asp	Val	Thr	
				710					715					720			

GenBank: C44471.60

	GCA	CCA	ACG	GTT	ACA	AAA	GTA	TTC	AAA	GGT	GAT	TCT	TTA	AAA	GAC	GCT	2304
	Ala	Pro	Thr	Val	Thr	Lys	Val	Phe	Lys	Gly	Asp	Ser	Leu	Lys	Asp	Ala	
			725					730					735				
5	GAT	GCA	GTT	ACT	ACA	CTT	ACG	AAC	GTT	GAT	GCA	GGT	CAA	AAA	TTC	ACT	2352
	Asp	Ala	Val	Thr	Thr	Leu	Thr	Asn	Val	Asp	Ala	Gly	Gln	Lys	Phe	Thr	
		740					745					750					
10	ATC	CAA	TTT	AGC	GAA	GAA	TTA	AAA	ACT	TCT	AGT	GGT	TCT	TTA	GTG	GGT	2400
	Ile	Gln	Phe	Ser	Glu	Glu	Leu	Lys	Thr	Ser	Ser	Gly	Ser	Leu	Val	Gly	
		755				760					765				770		
15	GGC	AAA	GTA	ACT	GTC	GAG	AAA	TTA	ACA	AAC	AAC	GGA	TGG	GTA	GAT	GCT	2448
	Gly	Lys	Val	Thr	Val	Glu	Lys	Leu	Thr	Asn	Asn	Gly	Trp	Val	Asp	Ala	
				775						780					785		
20	GGT	ACT	GGA	ACA	ACT	GTA	TCA	GTT	GCT	CCT	AAG	ACA	GAT	GCA	AAT	GGT	2496
	Gly	Thr	Gly	Thr	Thr	Val	Ser	Val	Ala	Pro	Lys	Thr	Asp	Ala	Asn	Gly	
				790					795					800			
25	AAA	GTA	ACA	GCT	GCT	GTG	GTT	ACA	TTA	ACT	GGT	CTT	GAC	AAT	AAC	GAC	2544
	Lys	Val	Thr	Ala	Ala	Val	Val	Thr	Leu	Thr	Gly	Leu	Asp	Asn	Asn	Asp	
			805					810					815				
30	AAA	GAT	GCG	AAA	TTG	CGT	CTG	GTA	GTA	GAT	AAG	TCT	TCT	ACT	GAT	GGA	2592
	Lys	Asp	Ala	Lys	Leu	Arg	Leu	Val	Val	Asp	Lys	Ser	Ser	Thr	Asp	Gly	
		820					825					830					
35	ATT	GCT	GAT	GTA	GCT	GGT	AAT	GTA	ATT	AAG	GAA	AAA	GAT	ATT	TTA	ATT	2640
	Ile	Ala	Asp	Val	Ala	Gly	Asn	Val	Ile	Lys	Glu	Lys	Asp	Ile	Leu	Ile	
		835				840					845				850		
40	CGT	TAC	AAC	AGC	TGG	AGA	CAC	ACT	GTA	GCT	TCT	GTG	AAA	GCT	GCT	GCT	2688
	Arg	Tyr	Asn	Ser	Trp	Arg	His	Thr	Val	Ala	Ser	Val	Lys	Ala	Ala	Ala	
				855						860					865		
45	GAC	AAA	GAT	GGT	CAA	AAC	GCT	TCT	GCT	GCA	TTC	CCA	ACA	AGC	ACT	GCA	2736
	Asp	Lys	Asp	Gly	Gln	Asn	Ala	Ser	Ala	Ala	Phe	Pro	Thr	Ser	Thr	Ala	
			870						875					880			
50	ATT	GAT	ACA	ACT	AAG	AGC	TTA	TTA	GTT	GAA	TTC	AAT	GAA	ACT	GAT	TTA	2784
	Ile	Asp	Thr	Thr	Lys	Ser	Leu	Leu	Val	Glu	Phe	Asn	Glu	Thr	Asp	Leu	
		885					890						895				
55	GCG	GAA	GTT	AAA	CCT	GAG	AAC	ATC	GTT	GTT	AAA	GAT	GCA	GCA	GGT	AAT	2832
	Ala	Glu	Val	Lys	Pro	Glu	Asn	Ile	Val	Val	Lys	Asp	Ala	Ala	Gly	Asn	
		900					905					910					
60	GCG	GTA	GCT	GGT	ACT	GTA	ACA	GCA	TTA	GAC	GGT	TCT	ACA	AAT	AAA	TTT	2880
	Ala	Val	Ala	Gly	Thr	Val	Thr	Ala	Leu	Asp	Gly	Ser	Thr	Asn	Lys	Phe	
		915				920					925					930	
65	GTA	TTC	ACT	CCA	TCT	CAA	GAA	TTA	AAA	GCT	GGT	ACA	GTT	TAC	TCT	GTA	2928
	Val	Phe	Thr	Pro	Ser	Gln	Glu	Leu	Lys	Ala	Gly	Thr	Val	Tyr	Ser	Val	
				935						940					945		
70	ACA	ATT	GAC	GGT	GTG	AGA	GAT	AAA	GTA	GGT	AAC	ACA	ATC	TCT	AAA	TAC	2976
	Thr	Ile	Asp	Gly	Val	Arg	Asp	Lys	Val	Gly	Asn	Thr	Ile	Ser	Lys	Tyr	
			950						955					960			
75	ATT	ACT	TCG	TTC	AAG	ACT	GTA	TCT	GCG	AAT	CCA	ACG	TTA	TCT	TCA	ATC	3024
	Ile	Thr	Ser	Phe	Lys	Thr	Val	Ser	Ala	Asn	Pro	Thr	Leu	Ser	Ser	Ile	
			965					970					975				
80	AGC	ATT	GCT	GAC	GGT	GCA	GTT	AAC	GTT	GAC	CGT	TCT	AAA	ACA	ATT	ACA	3072
	Ser	Ile	Ala	Asp	Gly	Ala	Val	Asn	Val	Asp	Arg	Ser	Lys	Thr	Ile	Thr	
		980					985					990					

00444-1000

	ATT GAA TTC AGC GAT TCA GTT CCA AAC CCA ACA ATC ACT CTT AAG AAG	3120
	Ile Glu Phe Ser Asp Ser Val Pro Asn Pro Thr Ile Thr Leu Lys Lys	
	995 1000 1005 1010	
5	GCT GAC GGA ACT TCA TTT ACT AAT TAC ACT TTA GTA AAT GTA AAT AAT	3168
	Ala Asp Gly Thr Ser Phe Thr Asn Tyr Thr Leu Val Asn Val Asn Asn	
	1015 1020 1025	
10	GAA AAT AAA ACA TAC AAA ATT GTA TTC CAC AAA GGT GTA ACA CTT GAC	3216
	Glu Asn Lys Thr Tyr Lys Ile Val Phe His Lys Gly Val Thr Leu Asp	
	1030 1035 1040	
15	GAG TTT ACT CAA TAT GAG TTA GCA GTT TCA AAA GAT TTT CAA ACT GGT	3264
	Glu Phe Thr Gln Tyr Glu Leu Ala Val Ser Lys Asp Phe Gln Thr Gly	
	1045 1050 1055	
	ACT GAT ATT GAT AGC AAA GTT ACA TTC ATC ACA GGT TCT GTT GCT ACT	3312
	Thr Asp Ile Asp Ser Lys Val Thr Phe Ile Thr Gly Ser Val Ala Thr	
	1060 1065 1070	
20	GAC GAA GTA AAA CCT GCT CTA GTA GGC GTT GGT TCA TGG AAT GGA ACA	3360
	Asp Glu Val Lys Pro Ala Leu Val Gly Val Gly Ser Trp Asn Gly Thr	
	1075 1080 1085 1090	
25	AGC TAT ACT CAG GAT GCT GCA GCA ACA CGA CTT CGG TCT GTA GCT GAC	3408
	Ser Tyr Thr Gln Asp Ala Ala Ala Thr Arg Leu Arg Ser Val Ala Asp	
	1095 1100 1105	
30	TTC GTT GCG GAG CCA GTT GCC CTT CAA TTC TCA GAA GGT ATC GAT TTA	3456
	Phe Val Ala Glu Pro Val Ala Leu Gln Phe Ser Glu Gly Ile Asp Leu	
	1110 1115 1120	
	ACG AAT GCA ACT GTG ACA GTA ACA AAT ATT ACT GAT GAT AAA ACT GTT	3504
	Thr Asn Ala Thr Val Thr Val Thr Asn Ile Thr Asp Asp Lys Thr Val	
35	1125 1130 1135	
	GAA GTT ATT TCA AAA GAG AGT GTA GAC GCA GAC CAT GAT GCA GGT GCT	3552
	Glu Val Ile Ser Lys Glu Ser Val Asp Ala Asp His Asp Ala Gly Ala	
	1140 1145 1150	
40	ACT AAG GAG ACA TTA GTA ATT AAC ACA GTT ACT CCT TTA GTA CTT GAT	3600
	Thr Lys Glu Thr Leu Val Ile Asn Thr Val Thr Pro Leu Val Leu Asp	
	1155 1160 1165 1170	
45	AAC AGC AAG ACT TAT AAG ATT GTT GTA AGT GGA GTT AAA GAT GCA GCA	3648
	Asn Ser Lys Thr Tyr Lys Ile Val Val Ser Gly Val Lys Asp Ala Ala	
	1175 1180 1185	
50	GGT AAT GTT GCA GAT ACT ATT ACA TTC TAT ATT AAG TAA	3687
	Gly Asn Val Ala Asp Thr Ile Thr Phe Tyr Ile Lys	
	1190 1195	

55 (2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1228 amino acids

(B) TYPE: amino acid

60 (D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

	Met	Asp	Arg	Lys	Lys	Ala	Val	Lys	Leu	Ala	Thr	Ala	Ser	Ala	Ile	Ala
	-30					-25					-20					-15
5	Ala	Ser	Ala	Phe	Val	Ala	Ala	Asn	Pro	Asn	Ala	Ser	Glu	Ala	Ala	Thr
				-10						-5					1	
	Asp	Val	Ala	Thr	Val	Val	Ser	Gln	Ala	Lys	Ala	Gln	Phe	Lys	Lys	Ala
			5					10					15			
10	Tyr	Tyr	Thr	Tyr	Ser	His	Thr	Val	Thr	Glu	Thr	Gly	Glu	Phe	Pro	Asn
		20					25					30				
	Ile	Asn	Asp	Val	Tyr	Ala	Glu	Tyr	Asn	Lys	Ala	Lys	Lys	Arg	Tyr	Arg
	35					40					45					50
	Asp	Ala	Val	Ala	Leu	Val	Asn	Lys	Ala	Gly	Gly	Ala	Lys	Lys	Asp	Ala
					55					60					65	
20	Tyr	Leu	Ala	Asp	Leu	Gln	Lys	Glu	Tyr	Glu	Thr	Tyr	Val	Phe	Lys	Ala
				70					75					80		
	Asn	Pro	Lys	Ser	Gly	Glu	Ala	Arg	Val	Ala	Thr	Tyr	Ile	Asp	Ala	Tyr
			85					90					95			
25	Asn	Tyr	Ala	Thr	Lys	Leu	Asp	Glu	Met	Arg	Gln	Glu	Leu	Glu	Ala	Ala
		100					105					110				
	Val	Gln	Ala	Lys	Asp	Leu	Glu	Lys	Ala	Glu	Gln	Tyr	Tyr	His	Lys	Ile
	115					120					125					130
30	Pro	Tyr	Glu	Ile	Lys	Thr	Arg	Thr	Val	Ile	Leu	Asp	Arg	Val	Tyr	Gly
					135					140					145	
35	Lys	Thr	Thr	Arg	Asp	Leu	Leu	Arg	Ser	Thr	Phe	Lys	Ala	Lys	Ala	Gln
				150					155					160		
	Glu	Leu	Arg	Asp	Ser	Leu	Ile	Tyr	Asp	Ile	Thr	Val	Ala	Met	Lys	Ala
			165					170					175			
40	Arg	Glu	Val	Gln	Asp	Ala	Val	Lys	Ala	Gly	Asn	Leu	Asp	Lys	Ala	Lys
		180					185					190				
	Ala	Ala	Val	Asp	Gln	Ile	Asn	Gln	Tyr	Leu	Pro	Lys	Val	Thr	Asp	Ala
	195					200					205					210
	Phe	Lys	Thr	Glu	Leu	Thr	Glu	Val	Ala	Lys	Lys	Ala	Leu	Asp	Ala	Asp
					215					220					225	
50	Glu	Ala	Ala	Leu	Thr	Pro	Lys	Val	Glu	Ser	Val	Ser	Ala	Ile	Asn	Thr
				230					235					240		
	Gln	Asn	Lys	Ala	Val	Glu	Leu	Thr	Ala	Val	Pro	Val	Asn	Gly	Thr	Leu
			245					250					255			
55	Lys	Leu	Gln	Leu	Ser	Ala	Ala	Ala	Asn	Glu	Asp	Thr	Val	Asn	Val	Asn
		260					265					270				
	Thr	Val	Arg	Ile	Tyr	Lys	Val	Asp	Gly	Asn	Ile	Pro	Phe	Ala	Leu	Asn
	275					280					285					290
	Thr	Ala	Asp	Val	Ser</											

	Gly	Ile	Lys	Asp	Lys	Asn	Gly	Lys	Glu	Phe	Lys	Glu	Asp	Ala	Phe	Thr	
			325					330					335				
5	Phe	Lys	Leu	Arg	Asn	Asp	Ala	Val	Val	Thr	Gln	Val	Phe	Gly	Thr	Asn	
		340					345					350					
	Val	Thr	Asn	Asn	Thr	Ser	Val	Asn	Leu	Ala	Ala	Gly	Thr	Phe	Asp	Thr	
		355				360					365					370	
10	Asp	Asp	Thr	Leu	Thr	Val	Val	Phe	Asp	Lys	Leu	Leu	Ala	Pro	Glu	Thr	
					375					380					385		
	Val	Asn	Ser	Ser	Asn	Val	Thr	Ile	Thr	Asp	Val	Glu	Thr	Gly	Lys	Arg	
				390					395					400			
15	Ile	Pro	Val	Ile	Ala	Ser	Thr	Ser	Gly	Ser	Thr	Ile	Thr	Ile	Thr	Leu	
			405					410					415				
20	Lys	Glu	Ala	Leu	Val	Thr	Gly	Lys	Gln	Tyr	Lys	Leu	Ala	Ile	Asn	Asn	
		420					425					430					
	Val	Lys	Thr	Leu	Thr	Gly	Tyr	Asn	Ala	Glu	Ala	Tyr	Glu	Leu	Val	Phe	
		435				440					445					450	
25	Thr	Ala	Asn	Ala	Ser	Ala	Pro	Thr	Val	Ala	Thr	Ala	Pro	Thr	Thr	Leu	
					455					460						465	
	Gly	Gly	Thr	Thr	Leu	Ser	Thr	Gly	Ser	Leu	Thr	Thr	Asn	Val	Trp	Gly	
				470					475					480			
30	Lys	Leu	Ala	Gly	Gly	Val	Asn	Glu	Ala	Gly	Thr	Tyr	Tyr	Pro	Gly	Leu	
			485					490					495				
35	Gln	Phe	Thr	Thr	Thr	Phe	Ala	Thr	Lys	Leu	Asp	Glu	Ser	Thr	Leu	Ala	
		500					505					510					
	Asp	Asn	Phe	Val	Leu	Val	Glu	Lys	Glu	Ser	Gly	Thr	Val	Val	Ala	Ser	
		515				520					525					530	
40	Glu	Leu	Lys	Tyr	Asn	Ala	Asp	Ala	Lys	Met	Val	Thr	Leu	Val	Pro	Lys	
					535					540					545		
	Ala	Asp	Leu	Lys	Glu	Asn	Thr	Ile	Tyr	Gln	Ile	Lys	Ile	Lys	Lys	Gly	
				550					555					560			
45	Leu	Lys	Ser	Asp	Lys	Gly	Ile	Glu	Leu	Gly	Thr	Val	Asn	Glu	Lys	Thr	
			565					570					575				
50	Tyr	Glu	Phe	Lys	Thr	Gln	Asp	Leu	Thr	Ala	Pro	Thr	Val	Ile	Ser	Val	
		580					585					590					
	Thr	Ser	Lys	Asn	Gly	Asp	Ala	Gly	Leu	Lys	Val	Thr	Glu	Ala	Gln	Glu	
		595				600					605					610	
55	Phe	Thr	Val	Lys	Phe	Ser	Glu	Asn	Leu	Asn	Thr	Phe	Asn	Ala	Thr	Thr	
					615					620					625		
	Val	Ser	Gly	Ser	Thr	Ile	Thr	Tyr	Gly	Gln	Val	Ala	Val	Val	Lys	Ala	
				630					635					640			
60	Gly	Ala	Asn	Leu	Ser	Ala	Leu	Thr	Ala	Ser	Asp	Ile	Ile	Pro	Ala	Ser	
			645					650					655				
65	Val	Glu	Ala	Val	Thr	Gly	Gln	Asp	Gly	Thr	Tyr	Lys	Val	Lys	Val	Ala	
		660					665					670					

Ala Asn Gln Leu Glu Arg Asn Gln Gly Tyr Lys Leu Val Val Phe Gly
675 680 685 690

5 Lys Gly Ala Thr Ala Pro Val Lys Asp Ala Ala Asn Ala Asn Thr Leu
695 700 705

Ala Thr Asn Tyr Ile Tyr Thr Phe Thr Thr Glu Gly Gln Asp Val Thr
710 715 720

10 Ala Pro Thr Val Thr Lys Val Phe Lys Gly Asp Ser Leu Lys Asp Ala
725 730 735

Asp Ala Val Thr Thr Leu Thr Asn Val Asp Ala Gly Gln Lys Phe Thr
740 745 750

15 Ile Gln Phe Ser Glu Glu Leu Lys Thr Ser Ser Gly Ser Leu Val Gly
755 760 765 770

20 Gly Lys Val Thr Val Glu Lys Leu Thr Asn Asn Gly Trp Val Asp Ala
775 780 785

Gly Thr Gly Thr Thr Val Ser Val Ala Pro Lys Thr Asp Ala Asn Gly
790 795 800

25 Lys Val Thr Ala Ala Val Val Thr Leu Thr Gly Leu Asp Asn Asn Asp
805 810 815

Lys Asp Ala Lys Leu Arg Leu Val Val Asp Lys Ser Ser Thr Asp Gly
820 825 830

30 Ile Ala Asp Val Ala Gly Asn Val Ile Lys Glu Lys Asp Ile Leu Ile
835 840 845 850

35 Arg Tyr Asn Ser Trp Arg His Thr Val Ala Ser Val Lys Ala Ala Ala
855 860 865

Asp Lys Asp Gly Gln Asn Ala Ser Ala Ala Phe Pro Thr Ser Thr Ala
870 875 880

40 Ile Asp Thr Thr Lys Ser Leu Leu Val Glu Phe Asn Glu Thr Asp Leu
885 890 895

Ala Glu Val Lys Pro Glu Asn Ile Val Val Lys Asp Ala Ala Gly Asn
900 905 910

45 Ala Val Ala Gly Thr Val Thr Ala Leu Asp Gly Ser Thr Asn Lys Phe
915 920 925 930

Val Phe Thr Pro Ser Gln Glu Leu Lys Ala Gly Thr Val Tyr Ser Val
935 940 945

50 Thr Ile Asp Gly Val Arg Asp Lys Val Gly Asn Thr Ile Ser Lys Tyr
950 955 960

55 Ile Thr Ser Phe Lys Thr Val Ser Ala Asn Pro Thr Leu Ser Ser Ile
965 970 975

Ser Ile Ala Asp Gly Ala Val Asn Val Asp Arg Ser Lys Thr Ile Thr
980 985 990

60 Ile Glu Phe Ser Asp Ser Val Pro Asn Pro Thr Ile Thr Leu Lys Lys
995 1000 1005 1010

65 Ala Asp Gly Thr Ser Phe Thr Asn Tyr Thr Leu Val Asn Val Asn Asn
1015 1020 1025

Glu	Asn	Lys	Thr	Tyr	Lys	Ile	Val	Phe	His	Lys	Gly	Val	Thr	Leu	Asp	
			1030					1035						1040		
Glu	Phe	Thr	Gln	Tyr	Glu	Leu	Ala	Val	Ser	Lys	Asp	Phe	Gln	Thr	Gly	
			1045				1050					1055				
Thr	Asp	Ile	Asp	Ser	Lys	Val	Thr	Phe	Ile	Thr	Gly	Ser	Val	Ala	Thr	
	1060					1065					1070					
Asp	Glu	Val	Lys	Pro	Ala	Leu	Val	Gly	Val	Gly	Ser	Trp	Asn	Gly	Thr	
1075					1080					1085					1090	
Ser	Tyr	Thr	Gln	Asp	Ala	Ala	Ala	Thr	Arg	Leu	Arg	Ser	Val	Ala	Asp	
			1095						1100					1105		
Phe	Val	Ala	Glu	Pro	Val	Ala	Leu	Gln	Phe	Ser	Glu	Gly	Ile	Asp	Leu	
			1110					1115					1120			
Thr	Asn	Ala	Thr	Val	Thr	Val	Thr	Asn	Ile	Thr	Asp	Asp	Lys	Thr	Val	
			1125				1130					1135				
Glu	Val	Ile	Ser	Lys	Glu	Ser	Val	Asp	Ala	Asp	His	Asp	Ala	Gly	Ala	
	1140					1145					1150					
Thr	Lys	Glu	Thr	Leu	Val	Ile	Asn	Thr	Val	Thr	Pro	Leu	Val	Leu	Asp	
1155					1160					1165					1170	
Asn	Ser	Lys	Thr	Tyr	Lys	Ile	Val	Val	Ser	Gly	Val	Lys	Asp	Ala	Ala	
			1175						1180					1185		
Gly	Asn	Val	Ala	Asp	Thr	Ile	Thr	Phe	Tyr	Ile	Lys					
			1190					1195								

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 base pairs
(B) TYPE: nucleotide
(C) STRANDEDNESS: single strand
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

TTAATCGATT CTAGATGGAT AGGAAAAAAG CTG

33

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 37 base pairs
(B) TYPE: nucleotide
(C) STRANDEDNESS: single strand
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

ATACCCGGGG GTACGGATCC GATACAGATT TGAGCAA

37

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 2766 base pairs
- (B) TYPE: nucleotide
- (C) STRANDEDNESS: both
- (D) TOPOLOGY: linear

(i) INITIAL ORIGIN:

- (A) ORGANISM: *Bacillus stearothermophilus*
- (B) STRAIN: PV72

(vii) IMMEDIATE ORIGIN:

- (B) CLONE(S): sbsB

(ix) CHARACTERISTIC:

- (A) NAME/KEY: CDS
- (B) POSTION: 1..2763

(ix) CHARACTERISTIC:

- (A) NAME/KEY: sig_peptide
- (B) POSTION: 1..93

(ix) CHARACTERISTIC:

- (A) NAME/KEY: mat_peptide
- (B) POSTION: 94..2763

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

ATG GCT TAT CAA CCT AAG TCT TTT CGC AAG TTT GTT GCG ACA ACT GCA	48
Met Ala Tyr Gln Pro Lys Ser Phe Arg Lys Phe Val Ala Thr Thr Ala	
-31 -30 -25 -20	
ACA GCT GCC ATT GTA GCA TCT GCG GTA GCT CCT GTA GTA TCT GCA GCA	96
Thr Ala Ala Ile Val Ala Ser Ala Val Ala Pro Val Val Ser Ala Ala	
-15 -10 -5 1	
AGC TTC ACA GAT GTT GCG CCG CAA TAT AAA GAT GCG ATC GAT TTC TTA	144
Ser Phe Thr Asp Val Ala Pro Gln Tyr Lys Asp Ala Ile Asp Phe Leu	
5 10 15	
GTA TCA ACT GGT GCA ACA AAA GGT AAA ACA GAA ACA AAA TTC GGC GTT	192
Val Ser Thr Gly Ala Thr Lys Gly Lys Thr Glu Thr Lys Phe Gly Val	
20 25 30	
TAC GAT GAA ATC ACT CGT CTA GAT GCG GCA GTT ATT CTT GCA AGA GTA	240
Tyr Asp Glu Ile Thr Arg Leu Asp Ala Ala Val Ile Leu Ala Arg Val	
35 40 45	

662023T 4444T60

	TTA	AAA	CTA	GAC	GTT	GAC	AAC	GCA	AAA	GAC	GCA	GGC	TTC	ACA	GAT	GTG	288
	Leu	Lys	Leu	Asp	Val	Asp	Asn	Ala	Lys	Asp	Ala	Gly	Phe	Thr	Asp	Val	
	50					55					60					65	
5	CCA	AAA	GAC	CGT	GCA	AAA	TAC	GTC	AAC	GCG	CTT	GTA	GAA	GCT	GGC	GTA	336
	Pro	Lys	Asp	Arg	Ala	Lys	Tyr	Val	Asn	Ala	Leu	Val	Glu	Ala	Gly	Val	
					70					75					80		
10	TTA	AAC	GGT	AAA	GCA	CCT	GGC	AAA	TTT	GGT	GCA	TAC	GAC	CCA	TTA	ACT	384
	Leu	Asn	Gly	Lys	Ala	Pro	Gly	Lys	Phe	Gly	Ala	Tyr	Asp	Pro	Leu	Thr	
				85					90					95			
15	CGC	GTT	GAA	ATG	GCA	AAA	ATC	ATC	GCG	AAC	CGT	TAC	AAA	TTA	AAA	GCT	432
	Arg	Val	Glu	Met	Ala	Lys	Ile	Ile	Ala	Asn	Arg	Tyr	Lys	Leu	Lys	Ala	
			100					105					110				
20	GAC	GAT	GTA	AAA	CTT	CCA	TTC	ACT	GAT	GTA	AAC	GAT	ACA	TGG	GCA	CCA	480
	Asp	Val	Lys	Leu	Pro	Phe	Thr	Asp	Val	Asn	Asp	Thr	Trp	Ala	Pro		
	115						120				125						
25	TAC	GTA	AAA	GCG	CTT	TAT	AAA	TAC	GAA	GTA	ACC	AAA	AGG	TTA	AAA	CAC	528
	Tyr	Val	Lys	Ala	Leu	Tyr	Lys	Tyr	Glu	Val	Thr	Lys	Arg	Leu	Lys	His	
	130					135					140					145	
30	CAA	CAA	GCT	TCG	GTG	CAT	ACC	AAA	AAC	ATC	ACT	CTG	CGT	GAC	TTT	GCG	576
	Gln	Gln	Ala	Ser	Val	His	Thr	Lys	Asn	Ile	Thr	Leu	Arg	Asp	Phe	Ala	
					150					155					160		
35	CAA	TTT	GTA	TAT	AGA	GCG	GTG	AAT	ATT	AAT	GCA	GTG	CCA	GAA	ATA	GTT	624
	Gln	Phe	Val	Tyr	Arg	Ala	Val	Asn	Ile	Asn	Ala	Val	Pro	Glu	Ile	Val	
				165					170					175			
40	GAA	GTA	ACT	GCG	GTT	AAT	TCG	ACT	ACA	GTG	AAA	GTA	ACA	TTC	AAT	ACG	672
	Glu	Val	Thr	Ala	Val	Asn	Ser	Thr	Thr	Val	Lys	Val	Thr	Phe	Asn	Thr	
			180					185					190				
45	CAA	ATT	GCT	GAT	GTT	GAT	TTC	ACA	AAT	TTT	GCT	ATC	GAT	AAC	GGT	TTA	720
	Gln	Ile	Ala	Asp	Val	Asp	Phe	Thr	Asn	Phe	Ala	Ile	Asp	Asn	Gly	Leu	
		195					200					205					
50	ACT	GTT	ACT	AAA	GCA	ACT	CTT	TCT	CGT	GAT	AAA	AAA	TCC	GTA	GAG	GTT	768
	Thr	Val	Thr	Lys	Ala	Thr	Leu	Ser	Arg	Asp	Lys	Lys	Ser	Val	Glu	Val	
	210					215					220					225	
55	GTG	GTA	AAT	AAA	CCG	TTT	ACT	CGT	AAT	CAG	GAA	TAT	ACA	ATT	ACA	GCG	816
	Val	Val	Asn	Lys	Pro	Phe	Thr	Arg	Asn	Gln	Glu	Tyr	Thr	Ile	Thr	Ala	
				230						235					240		
60	ACA	GGC	ATT	AAA	AAT	TTA	AAA	GGC	GAG	ACC	GCT	AAG	GAA	TTA	ACT	GGT	864
	Thr	Gly	Ile	Lys	Asn	Leu	Lys	Gly	Glu	Thr	Ala	Lys	Glu	Leu	Thr	Gly	
				245					250					255			
65	AAG	TTT	GTT	TGG	TCT	GTT	CAA	GAT	GCG	GTA	ACT	GTT	GCA	CTA	AAT	AAT	912
	Lys	Phe	Val	Trp	Ser	Val	Gln	Asp	Ala	Val	Thr	Val	Ala	Leu	Asn	Asn	
			260					265					270				
70	AGT	TCG	CTT	AAA	GTT	GGA	GAG	GAA	TCT	GGT	TTA	ACT	GTA	AAA	GAT	CAG	960
	Ser	Ser	Leu	Lys	Val	Gly	Glu	Glu	Ser	Gly	Leu	Thr	Val	Lys	Asp	Gln	
			275				280					285					
75	GAT	GGC	AAA	GAT	GTT	GTA	GGT	GCT	AAA	GTA	GAA	CTT	ACT	TCT	TCT	AAT	1008
	Asp	Gly	Lys	Asp	Val	Val	Gly	Ala	Lys	Val	Glu	Leu	Thr	Ser	Ser	Asn	
					290		295				300					305	
80	ACT	AAT	ATT	GTT	GTA	GTT	TCA	AGT	GGC	GAA	GTA	TCA	GTA	TCT	GCT	GCT	1056
	Thr	Asn	Ile	Val	Val	Val	Ser	Ser	Gly	Glu	Val	Ser	Val	Ser	Ala	Ala	
					310					315					320		

Genetic Code

	AAA	GTT	ACA	GCT	GTA	AAA	CCG	GGA	ACA	GCT	GAT	GTT	ACT	GCA	AAA	GTT	1104
	Lys	Val	Thr	Ala	Val	Lys	Pro	Gly	Thr	Ala	Asp	Val	Thr	Ala	Lys	Val	
				325					330					335			
5	ACA	TTA	CCA	GAT	GGT	GTT	GTA	CTA	ACA	AAT	ACA	TTT	AAA	GTG	ACA	GTT	1152
	Thr	Leu	Pro	Asp	Gly	Val	Val	Leu	Thr	Asn	Thr	Phe	Lys	Val	Thr	Val	
			340					345					350				
10	ACA	GAA	GTG	CCT	GTT	CAA	GTC	CAA	AAT	CAA	GGA	TTT	ACT	TTA	GTT	GAT	1200
	Thr	Glu	Val	Pro	Val	Gln	Val	Gln	Asn	Gln	Gly	Phe	Thr	Leu	Val	Asp	
		355					360					365					
15	AAT	CTT	TCT	AAT	GCT	CCA	CAG	AAT	ACA	GTT	GCA	TTT	AAC	AAA	GCT	GAG	1248
	Asn	Leu	Ser	Asn	Ala	Pro	Gln	Asn	Thr	Val	Ala	Phe	Asn	Lys	Ala	Glu	
	370					375					380					385	
20	AAA	GTA	ACT	TCA	ATG	TTT	GCT	GGA	GAA	ACT	AAA	ACA	GTT	GCA	ATG	TAT	1296
	Lys	Val	Thr	Ser	Met	Phe	Ala	Gly	Glu	Thr	Lys	Thr	Val	Ala	Met	Tyr	
					390					395					400		
25	GAT	ACT	AAA	AAC	GGT	GAT	CCT	GAA	ACT	AAA	CCT	GTT	GAT	TTC	AAA	GAT	1344
	Asp	Thr	Lys	Asn	Gly	Asp	Pro	Glu	Thr	Lys	Pro	Val	Asp	Phe	Lys	Asp	
				405					410					415			
30	GCA	ACT	GTA	CGT	TCA	TTA	AAT	CCA	ATT	ATT	GCA	ACA	GCT	GCT	ATT	AAT	1392
	Ala	Thr	Val	Arg	Ser	Leu	Asn	Pro	Ile	Ile	Ala	Thr	Ala	Ala	Ile	Asn	
			420					425					430				
35	GGT	AGT	GAG	CTC	CTT	GTC	ACA	GCT	AAT	GCT	GGC	CAA	TCT	GGA	AAA	GCT	1440
	Gly	Ser	Glu	Leu	Leu	Val	Thr	Ala	Asn	Ala	Gly	Gln	Ser	Gly	Lys	Ala	
		435					440					445					
40	TCA	TTT	GAA	GTA	ACA	TTA	AAA	GAT	AAT	ACA	AAA	AGA	ACA	TTT	ACA	GTT	1488
	Ser	Phe	Glu	Val	Thr	Leu	Lys	Asp	Asn	Thr	Lys	Arg	Thr	Phe	Thr	Val	
	450					455					460					465	
45	GAT	GTA	AAA	AAA	GAC	CCT	GTA	TTA	CAA	GAT	ATA	AAA	GTA	GAT	GCA	ACT	1536
	Asp	Val	Lys	Lys	Asp	Pro	Val	Leu	Gln	Asp	Ile	Lys	Val	Asp	Ala	Thr	
					470					475					480		
50	TCT	GTT	AAA	CTT	TCC	GAT	GAA	GCT	GTT	GGC	GGC	GGG	GAA	GTT	GAA	GGA	1584
	Ser	Val	Lys	Leu	Ser	Asp	Glu	Ala	Val	Gly	Gly	Gly	Glu	Val	Glu	Gly	
				485				490					495				
55	GTT	AAC	CAA	AAA	ACG	ATT	AAA	GTA	AGT	GCA	GTT	GAC	CAA	TAC	GGT	AAA	1632
	Val	Asn	Gln	Lys	Thr	Ile	Lys	Val	Ser	Ala	Val	Asp	Gln	Tyr	Gly	Lys	
			500					505					510				
60	GAA	ATT	AAA	TTT	GGT	ACA	AAA	GGT	AAA	GTT	ACT	GTT	ACA	ACT	AAT	ACA	1680
	Glu	Ile	Lys	Phe	Gly	Thr	Lys	Gly	Lys</								

	AAT	GTA	AAA	GCT	GAC	GCT	ACA	CCA	GTA	GGA	TTA	GAT	ATT	GTA	GCA	CCT	1920
	Asn	Val	Lys	Ala	Asp	Ala	Thr	Pro	Val	Gly	Leu	Asp	Ile	Val	Ala	Pro	
		595					600					605					
5	TCT	AAA	ATT	GAT	GTA	AAT	GCT	CCA	AAC	ACT	GCT	TCT	ACT	GCA	GAT	GTT	1968
	Ser	Lys	Ile	Asp	Val	Asn	Ala	Pro	Asn	Thr	Ala	Ser	Thr	Ala	Asp	Val	
	610					615					620					625	
	GAT	TTT	ATA	AAT	TTC	GAA	AGT	GTT	GAG	ATT	TAC	ACA	CTC	GAT	TCA	AAT	2016
10	Asp	Phe	Ile	Asn	Phe	Glu	Ser	Val	Glu	Ile	Tyr	Thr	Leu	Asp	Ser	Asn	
					630					635						640	
	GGT	AGA	CGT	CAA	AAA	AAA	GTT	ACT	CCA	ACT	GCA	ACT	ACA	CTT	GTA	GGT	2064
15	Gly	Arg	Arg	Gln	Lys	Lys	Val	Thr	Pro	Thr	Ala	Thr	Thr	Leu	Val	Gly	
				645					650					655			
	ACA	AAA	AAA	AAA	AAA	AAA	GTT	AAT	GGG	AAT	GTA	TTA	CAA	TTC	AAG	GGG	2112
	Thr	Lys	Lys	Lys	Lys	Lys	Val	Asn	Gly	Asn	Val	Leu	Gln	Phe	Lys	Gly	
				660				665					670				
20	AAC	GAA	GAA	TTA	ACG	CTA	TCA	ACT	TCT	TCT	AGT	ACA	GGA	AAC	GTA	GAT	2160
	Asn	Glu	Glu	Leu	Thr	Leu	Ser	Thr	Ser	Ser	Ser	Thr	Gly	Asn	Val	Asp	
		675					680					685					
25	GGA	ACA	GCA	GAA	GGA	ATG	ACA	AAA	CGT	ATT	CCA	GGG	AAA	TAT	ATC	AAC	2208
	Gly	Thr	Ala	Glu	Gly	Met	Thr	Lys	Arg	Ile	Pro	Gly	Lys	Tyr	Ile	Asn	
	690					695					700					705	
	TCT	GCA	AGT	GTA	CCT	GCC	AGT	GCA	ACA	GTA	GCA	ACA	AGT	CCT	GTT	ACT	2256
30	Ser	Ala	Ser	Val	Pro	Ala	Ser	Ala	Thr	Val	Ala	Thr	Ser	Pro	Val	Thr	
					710					715					720		
	GTA	AAG	CTT	AAT	TCA	AGT	GAT	AAT	GAT	TTA	ACA	TTT	GAA	GAA	TTA	ATA	2304
35	Val	Lys	Leu	Asn	Ser	Ser	Asp	Asn	Asp	Leu	Thr	Phe	Glu	Glu	Leu	Ile	
				725					730					735			
	TTC	GGT	GTA	ATT	GAC	CCT	ACA	CAA	TTA	GTC	AAA	GAT	GAA	GAC	ATC	AAC	2352
	Phe	Gly	Val	Ile	Asp	Pro	Thr	Gln	Leu	Val	Lys	Asp	Glu	Asp	Ile	Asn	
			740					745					750				
40	GAA	TTT	ATT	GCA	GTT	TCA	AAA	GCG	GCT	AAA	AAT	GAT	GGA	TAT	TTG	TAT	2400
	Glu	Phe	Ile	Ala	Val	Ser	Lys	Ala	Lys	Asn	Asp	Gly	Tyr	Leu	Tyr		
		755					760				765						
45	AAT	AAA	CCG	CTT	GTA	ACG	GTT	AAA	GAT	GCA	TCA	GGA	AAA	GTT	ATT	CCA	2448
	Asn	Lys	Pro	Leu	Val	Thr	Val	Lys	Asp	Ala	Ser	Gly	Lys	Val	Ile	Pro	
	770					775					780					785	
	ACA	GGT	GCA	AAT	GTT	TAC	GGT	CTA	AAT	CAT	GAT	GCA	ACT	AAC	GGA	AAC	2496
50	Thr	Gly	Ala	Asn	Val	Tyr	Gly	Leu	Asn	His	Asp	Ala	Thr	Asn	Gly	Asn	
					790					795					800		
	ATT	TGG	TTT	GAT	GAG	GAA	CAA	GCT	GGC	TTA	GCT	AAA	AAA	TTT	AGT	GAT	2544
55	Ile	Trp	Phe	Asp	Glu	Glu	Gln	Ala	Gly	Leu	Ala	Lys	Lys	Phe	Ser	Asp	
				805					810					815			
	GTA	CAT	TTT	GAT	GTT	GAT	TTT	TCA	TTA	ACT	AAC	GTT	GTA	AAA	ACT	GGT	2592
	Val	His	Phe	Asp	Val	Asp	Phe	Ser	Leu	Thr	Asn	Val	Val	Lys	Thr	Gly	
			820				825						830				
60	AGC	GGT	ACA	GTT	TCT	TCA	TCG	CCA	TCA	TTA	TCT	GAC	GCA	ATT	CAA	CTT	2640
	Ser	Gly	Thr	Val	Ser	Ser	Ser	Pro	Ser	Leu	Ser	Asp	Ala	Ile	Gln	Leu	
		835					840					845					
65	ACT	AAT	TCA	GGC	GAT	GCA	GTA	TCG	TTT	ACA	TTA	GTT	ATC	AAA	TCA	ATT	2688
	Thr	Asn	Ser	Gly	Asp	Ala	Val	Ser	Phe	Thr	Leu	Val	Ile	Lys	Ser	Ile	
	850					855					860					865	

GenBank: 241272

TAT GTT AAA GGC GCA GAT AAA GAT GAT AAT AAC TTA CTT GCA GCC CCT 2736
Tyr Val Lys Gly Ala Asp Lys Asp Asp Asn Asn Leu Leu Ala Ala Pro
870 875 880

5 GTT TCT GTC AAT GTG ACT GTG ACA AAA TAA 2766
Val Ser Val Asn Val Thr Val Thr Lys
885 890

10 (2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 921 amino acids
(B) TYPE: amino acid
15 (D) TOPOLOGY: linear

(ii) TYPE OF MOLECULE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

20

Met Ala Tyr Gln Pro Lys Ser Phe Arg Lys Phe Val Ala Thr Thr Ala
25 -31 -30 . -25 -20
Thr Ala Ala Ile Val Ala Ser Ala Val Ala Pro Val Val Ser Ala Ala
-15 -10 -5 1
30 Ser Phe Thr Asp Val Ala Pro Gln Tyr Lys Asp Ala Ile Asp Phe Leu
5 10 15
Val Ser Thr Gly Ala Thr Lys Gly Lys Thr Glu Thr Lys Phe Gly Val
20 25 30
35 Tyr Asp Glu Ile Thr Arg Leu Asp Ala Ala Val Ile Leu Ala Arg Val
35 40 45
Leu Lys Leu Asp Val Asp Asn Ala Lys Asp Ala Gly Phe Thr Asp Val
40 50 55 60 65
Pro Lys Asp Arg Ala Lys Tyr Val Asn Ala Leu Val Glu Ala Gly Val
70 75 80
45 Leu Asn Gly Lys Ala Pro Gly Lys Phe Gly Ala Tyr Asp Pro Leu Thr
85 90 95
Arg Val Glu Met Ala Lys Ile Ile Ala Asn Arg Tyr Lys Leu Lys Ala
100 105 110
50 Asp Asp Val Lys Leu Pro Phe Thr Asp Val Asn Asp Thr Trp Ala Pro
115 120 125
Tyr Val Lys Ala Leu Tyr Lys Tyr Glu Val Thr Lys Arg Leu Lys His
55 130 135 140 145
Gln Gln Ala Ser Val His Thr Lys Asn Ile Thr Leu Arg Asp Phe Ala
150 155 160
60 Gln Phe Val Tyr Arg Ala Val Asn Ile Asn Ala Val Pro Glu Ile Val
165 170 175
Glu Val Thr Ala Val Asn Ser Thr Thr Val Lys Val Thr Phe Asn Thr
180 185 190

	Gln	Ile	Ala	Asp	Val	Asp	Phe	Thr	Asn	Phe	Ala	Ile	Asp	Asn	Gly	Leu
	195					200						205				
5	Thr	Val	Thr	Lys	Ala	Thr	Leu	Ser	Arg	Asp	Lys	Lys	Ser	Val	Glu	Val
	210					215					220					225
	Val	Val	Asn	Lys	Pro	Phe	Thr	Arg	Asn	Gln	Glu	Tyr	Thr	Ile	Thr	Ala
					230					235					240	
10	Thr	Gly	Ile	Lys	Asn	Leu	Lys	Gly	Glu	Thr	Ala	Lys	Glu	Leu	Thr	Gly
				245					250					255		
	Lys	Phe	Val	Trp	Ser	Val	Gln	Asp	Ala	Val	Thr	Val	Ala	Leu	Asn	Asn
			260					265					270			
15	Ser	Ser	Leu	Lys	Val	Gly	Glu	Glu	Ser	Gly	Leu	Thr	Val	Lys	Asp	Gln
		275					280					285				
20	Asp	Gly	Lys	Asp	Val	Val	Gly	Ala	Lys	Val	Glu	Leu	Thr	Ser	Ser	Asn
	290					295					300					305
	Thr	Asn	Ile	Val	Val	Val	Ser	Ser	Gly	Glu	Val	Ser	Val	Ser	Ala	Ala
					310					315					320	
25	Lys	Val	Thr	Ala	Val	Lys	Pro	Gly	Thr	Ala	Asp	Val	Thr	Ala	Lys	Val
				325					330					335		
	Thr	Leu	Pro	Asp	Gly	Val	Val	Leu	Thr	Asn	Thr	Phe	Lys	Val	Thr	Val
			340					345					350			
30	Thr	Glu	Val	Pro	Val	Gln	Val	Gln	Asn	Gln	Gly	Phe	Thr	Leu	Val	Asp
		355					360					365				
35	Asn	Leu	Ser	Asn	Ala	Pro	Gln	Asn	Thr	Val	Ala	Phe	Asn	Lys	Ala	Glu
	370					375					380					385
	Lys	Val	Thr	Ser	Met	Phe	Ala	Gly	Glu	Thr	Lys	Thr	Val	Ala	Met	Tyr
					390					395					400	
40	Asp	Thr	Lys	Asn	Gly	Asp	Pro	Glu	Thr	Lys	Pro	Val	Asp	Phe	Lys	Asp
				405					410					415		
	Ala	Thr	Val	Arg	Ser	Leu	Asn	Pro	Ile	Ile	Ala	Thr	Ala	Ala	Ile	Asn
			420					425					430			
45	Gly	Ser	Glu	Leu	Leu	Val	Thr	Ala	Asn	Ala	Gly	Gln	Ser	Gly	Lys	Ala
		435					440					445				
50	Ser	Phe	Glu	Val	Thr	Leu	Lys	Asp	Asn	Thr	Lys	Arg	Thr	Phe	Thr	Val
	450					455					460					465
	Asp	Val	Lys	Lys	Asp	Pro	Val	Leu	Gln	Asp	Ile	Lys	Val	Asp	Ala	Thr
					470					475					480	
55	Ser	Val	Lys	Leu	Ser	Asp	Glu	Ala	Val	Gly	Gly	Gly	Glu	Val	Glu	Gly
				485					490					495		
	Val	Asn	Gln	Lys	Thr	Ile	Lys	Val	Ser	Ala	Val	Asp	Gln	Tyr	Gly	Lys
			500					505					510			
60	Glu	Ile	Lys	Phe	Gly	Thr	Lys	Gly	Lys	Val	Thr	Val	Thr	Thr	Asn	Thr
		515					520					525				

	Asp	Ser	Gly	Asn	Ser	Ala	Thr	Asp	Gln	Phe	Val	Val	Val	Ala	Thr	Lys
				550						555					560	
5	Asp	Lys	Ile	Val	Asn	Gly	Lys	Val	Glu	Val	Lys	Tyr	Phe	Lys	Asn	Ala
			565						570					575		
	Ser	Asp	Thr	Thr	Pro	Thr	Ser	Thr	Lys	Thr	Ile	Thr	Val	Asn	Val	Val
			580					585					590			
10	Asn	Val	Lys	Ala	Asp	Ala	Thr	Pro	Val	Gly	Leu	Asp	Ile	Val	Ala	Pro
		595					600					605				
	Ser	Lys	Ile	Asp	Val	Asn	Ala	Pro	Asn	Thr	Ala	Ser	Thr	Ala	Asp	Val
	610					615					620				625	
15	Asp	Phe	Ile	Asn	Phe	Glu	Ser	Val	Glu	Ile	Tyr	Thr	Leu	Asp	Ser	Asn
				630						635					640	
	Gly	Arg	Arg	Gln	Lys	Lys	Val	Thr	Pro	Thr	Ala	Thr	Thr	Leu	Val	Gly
20				645					650					655		
	Thr	Lys	Lys	Lys	Lys	Lys	Val	Asn	Gly	Asn	Val	Leu	Gln	Phe	Lys	Gly
			660					665					670			
25	Asn	Glu	Glu	Leu	Thr	Leu	Ser	Thr	Ser	Ser	Ser	Thr	Gly	Asn	Val	Asp
		675					680					685				
	Gly	Thr	Ala	Glu	Gly	Met	Thr	Lys	Arg	Ile	Pro	Gly	Lys	Tyr	Ile	Asn
	690					695					700				705	
30	Ser	Ala	Ser	Val	Pro	Ala	Ser	Ala	Thr	Val	Ala	Thr	Ser	Pro	Val	Thr
					710					715					720	
	Val	Lys	Leu	Asn	Ser	Ser	Asp	Asn	Asp	Leu	Thr	Phe	Glu	Glu	Leu	Ile
35				725					730					735		
	Phe	Gly	Val	Ile	Asp	Pro	Thr	Gln	Leu	Val	Lys	Asp	Glu	Asp	Ile	Asn
			740					745					750			
40	Glu	Phe	Ile	Ala	Val	Ser	Lys	Ala	Ala	Lys	Asn	Asp	Gly	Tyr	Leu	Tyr
		755					760					765				
	Asn	Lys	Pro	Leu	Val	Thr	Val	Lys	Asp	Ala	Ser	Gly	Lys	Val	Ile	Pro
	770					775					780				785	
45	Thr	Gly	Ala	Asn	Val	Tyr	Gly	Leu	Asn	His	Asp	Ala	Thr	Asn	Gly	Asn
				790						795					800	
	Ile	Trp	Phe	Asp	Glu	Glu	Gln	Ala	Gly	Leu	Ala	Lys	Lys	Phe	Ser	Asp
50			805						810					815		
	Val	His	Phe	Asp	Val	Asp	Phe	Ser	Leu	Thr	Asn	Val	Val	Lys	Thr	Gly
			820					825					830			
55	Ser	Gly	Thr	Val	Ser	Ser	Ser	Pro	Ser	Leu	Ser	Asp	Ala	Ile	Gln	Leu
		835					840					845				
	Thr	Asn	Ser	Gly	Asp	Ala	Val	Ser	Phe	Thr	Leu	Val	Ile	Lys	Ser	Ile
	850					855					860				865	
60	Tyr	Val	Lys	Gly	Ala	Asp	Lys	Asp	Asp	Asn	Asn	Leu	Leu	Ala	Ala	Pro
				870						875					880	
	Val	Ser	Val	Asn	Val	Thr	Val	Thr	Lys							
65				885					890							

(2) INFORMATION FOR SEQ ID NO: 7:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 498 base pairs
 - (B) TYPE: nucleotide
 - (C) STRANDEDNESS: both
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

CCCATGGACC CGTCCAAGGA CTCCAAAGCT CAGGTTTCTG CAGCCGAAGC TGGTATCACT	60
GGCACCTGGT ATAACCAACT GGGGTCGACT TTCATTGTGA CCGCTGGTGC GGACGGAGCT	120
CTGACTGGCA CCTACGAATC TGCGGTTGGT AACGCAGAAT CCCGCTACGT ACTGACTGGC	180
CGTTATGACT CTGCACCTGC CACCGATGGC TCTGGTACCG CTCTGGGCTG GACTGTGGCT	240
TGGAAAAACA ACTATCGTAA TGCGCACAGC GCCACTACGT GGTCTGGCCA ATACGTTGGC	300
GGTGCTGAGG CTCGTATCAA CACTCAGTGG CTGTTAACAT CCGGCACTAC CGAAGCGAAT	360
GCATGGAAAT CGACACTAGT AGGTCATGAC ACCTTTACCA AAGTTAAGCC TTCTGCTGCT	420
AGCATTGATG CTGCCAAGAA AGCAGGCGTA AACACGGTA ACCCTCTAGA CGCTGTTTCAG	480
CAATAATAAG GATCCGGG	498

(2) INFORMATION FOR SEQ ID NO: 8:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 29 base pairs
 - (B) TYPE: nucleotide
 - (C) STRANDEDNESS: single strand
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8

TTCATCGTAA ACGCCGAATT TTGTTTCTG	29
---------------------------------	----

INFORMATION FOR SEQ ID NO: 9:

- (i) SEQUENCE CHARACTERISTICS:
 - (A) LENGTH: 26 base pairs
 - (B) TYPE: nucleotide
 - (C) STRANDEDNESS: single strand
 - (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:

AGGGAAATAT ATCAACTCTG CAAGTG	26
------------------------------	----

Claims

1. Process for the production of S-layer proteins
w h e r e i n
 - (a) a gram-negative prokaryotic host cell is provided which is transformed with a nucleic acid coding for an S-layer protein which is selected from
 - (i) a nucleic acid which comprises the nucleotide sequence from position 1 to 3684 shown in SEQ ID NO.1 optionally without the signal peptide-coding section,
 - (ii) a nucleic acid which comprises a nucleotide sequence corresponding to the nucleic acid from (i) within the scope of the degeneracy of the genetic code and
 - (iii) a nucleic acid which comprises a nucleotide sequence which hybridizes with the nucleic acids from (i) or/and (ii) under stringent conditions;
 - (b) the host cell is cultured under conditions which lead to an expression of the nucleic acid and to production of the polypeptide coded by it and
 - (c) the resulting polypeptide is isolated from the host cell.
2. Process as claimed in claim 1,
w h e r e i n
an E. coli host cell is used.

3. Process as claimed in claim 1 or 2,
w h e r e i n
the polypeptide is isolated from the interior of
the host cell in the form of an assembled S-layer
structure.
4. Process as claimed in one of the claims 1 to 3,
w h e r e i n
the nucleic acid coding for the S-layer protein
contains one or several insertions which code for
peptide or polypeptide sequences.
5. Process as claimed in claim 4,
w h e r e i n
the insertions are selected from nucleotide
sequences which code for cysteine residues, regions
with several charged amino acids or Tyr residues,
DNA-binding epitopes, metal-binding epitopes,
immunogenic epitopes, allergenic epitopes,
antigenic epitopes, streptavidin, enzymes,
cytokines or antibody-binding proteins.
6. Process as claimed in claim 5,
w h e r e i n
the insertions code for streptavidin.
7. Process as claimed in claim 5,
w h e r e i n
the insertions code for immunogenic epitopes from
herpes viruses, in particular herpes virus 6 or
FMDV.

8. Process as claimed in claim 5,
w h e r e i n
the insertions code for enzymes such as
polyhydroxybutyric acid synthase or bacterial
luciferase.
9. Process as claimed in claim 5,
w h e r e i n
the insertions code for cytokines such as
interleukins, interferons or tumour necrosis
factors.
10. Process as claimed in claim 5,
w h e r e i n
the insertions code for antibody-binding proteins
such as protein A or protein G.
11. Process as claimed in claim 5,
w h e r e i n
the insertions code for antigenic epitopes which
bind cytokines or endotoxins.
12. Process as claimed in claim 5,
w h e r e i n
the insertions code for metal-binding epitopes.
13. Process as claimed in one of the claims 1 to 12,
w h e r e i n
a nucleic acid coding for a gram-positive signal
peptide is arranged in operative linkage at the 5'
side of the nucleic acid coding for the S-layer
protein.

14. Process as claimed in claim 13,
w h e r e i n
the nucleic acid coding for the signal peptide
comprises
- (a) the signal peptide-coding section of the
nucleotide sequence shown in SEQ ID NO.1,
 - (b) a nucleotide sequence corresponding to the
sequence from (a) within the degeneracy of the
genetic code or/and
 - (c) a nucleotide sequence that is at least 80 %
homologous to the sequences from (a) or/and
(b).
15. Nucleic acid that codes for a recombinant S-layer
protein and is selected from
- (i) a nucleic acid which comprises the nucleotide
sequence from position 1 to 3684 shown in
SEQ ID NO.1 optionally without the signal
peptide-coding section,
 - (ii) a nucleic acid which comprises a nucleotide
sequence corresponding to the nucleic acid
from (i) within the scope of the degeneracy
of the genetic code and
 - (iii) a nucleic acid which comprises a nucleotide
sequence which hybridizes with one of the
nucleic acids from (i) or/and (ii) under
stringent conditions,
- wherein the nucleic acid contains at least one
peptide or polypeptide-coding insertion within the
region coding for the S-layer protein.
16. Nucleic acid as claimed in claim 15,
w h e r e i n
the insertion site is located at position 582, 878,

917, 2504 or/and 2649 of the nucleotide sequence shown in SEQ ID NO.1.

17. Vector,
w h e r e i n
it contains at least one copy of a nucleic acid as claimed in claim 15 or 16.
18. Cell,
w h e r e i n
it is transformed with a nucleic acid as claimed in claim 15 or 16 or with a vector as claimed in claim 17.
19. Cell as claimed in claim 18,
w h e r e i n
it is a gram-negative prokaryotic cell and in particular an E. coli cell.
20. Cell as claimed in claim 18 or 19,
w h e r e i n
it contains a recombinant S-layer structure.
21. Recombinant S-layer protein,
w h e r e i n
it is coded by a nucleic acid as claimed in claim 15 or 16.
22. Recombinant S-layer structure,
w h e r e i n
it contains at least one protein as claimed in claim 21 as a subunit.

23. S-layer structure as claimed in claim 22,
w h e r e i n
it additionally contains at least one unmodified S-
layer protein as a subunit.
24. S-layer structure as claimed in claim 22 or 23,
w h e r e i n
it comprises several layers which are linked
covalently or by affinity binding.
25. Use of an S-layer protein as claimed in claim 21 or
an S-layer structure as claimed in one of the claims
22 to 24 as a vaccine or adjuvant.
26. Use as claimed in claim 25,
w h e r e i n
the vaccine or adjuvant additionally comprise a
bacterial ghost which optionally contains further
immunogenic epitopes in its membrane.
27. Use of an S-layer protein as claimed in claim 21 or
an S-layer structure as claimed in one of the claims
22 to 24 as an enzyme reactor.
28. Nucleic acid which codes for an S-layer protein and
is selected from
 - (i) a nucleic acid which comprises the nucleotide
sequence from position 1 to 2763 shown in SEQ
ID NO.5 optionally without the signal peptide-
coding section,
 - (ii) a nucleic acid which comprises a nucleotide
sequence corresponding to the nucleic acid from
(i) within the scope of the degeneracy of the
genetic code and

(iii) a nucleic acid which comprises a nucleotide sequence that hybridizes with the nucleic acids from (i) or/and (ii) under stringent conditions.

29. Nucleic acid as claimed in claim 28,
w h e r e i n
it contains at least one peptide-coding or
polypeptide-coding insertion within the region
coding for the S-layer protein.
30. Vector,
w h e r e i n
it contains at least one copy of a nucleic acid as
claimed in claim 28 or 29.
31. Cell,
w h e r e i n
it is transformed with a nucleic acid as claimed in
claim 28 or 29 or with a vector as claimed in claim
30.
32. Cell as claimed in claim 31,
w h e r e i n
it contains a recombinant S-layer structure.
33. S-layer protein,
w h e r e i n
it is coded by a nucleic acid as claimed in claim
29.

34. Recombinant S-layer structure,
w h e r e i n
it contains at least one recombinant S-layer
protein as a subunit which is coded by a nucleic
acid as claimed in claim 29.
35. Use of an S-layer protein as claimed in claim 33 or
of an S-layer structure as claimed in claim 34 as a
vaccine or adjuvant.
36. Use of an S-layer protein as claimed in claim 33 or
an S-layer structure as claimed in claim 34 as an
enzyme reactor.
37. Process for the production of recombinant S-layer
proteins,
w h e r e i n
(a) a host cell is provided which contains a
nucleic acid coding for an S-layer protein
which contains a peptide-coding or polypeptide-
coding insertion within the region coding for
the S-layer protein,
(b) the host cell is cultured under conditions
which lead to an expression of the nucleic acid
and to production of the polypeptide coded by
it and
(c) the resulting polypeptide is isolated from the
host cell or from the culture medium.
38. Process as claimed in claim 37,
w h e r e i n
the nucleic acid coding for the recombinant S-layer
protein is selected from
(i) a nucleic acid which comprises the nucleotide

sequence from position 1 to 3684 shown in SEQ ID NO.1 optionally without the signal peptide-coding section,

- (ii) a nucleic acid which comprises a nucleotide sequence corresponding to the nucleic acid from (i) within the scope of the degeneracy of the genetic code and
- (iii) a nucleic acid which comprises a nucleotide sequence which hybridizes with one of the nucleic acids from (i) or/and (ii) under stringent conditions

39. Process as claimed in claim 37,

w h e r e i n

the nucleic acid which codes for the recombinant S-layer protein is selected from

- (i) a nucleic acid which comprises the nucleotide sequence from position 1 to 2763 shown in SEQ ID NO.5 optionally without the signal peptide-coding section,
- (ii) a nucleic acid which comprises a nucleotide sequence corresponding to the nucleic acid from (i) within the scope of the degeneracy of the genetic code and
- (iii) a nucleic acid which comprises a nucleotide sequence that hybridizes with the nucleic acids from (i) or/and (ii) under stringent conditions.

40. Process as claimed in one of the claims 37-39,

w h e r e i n

a further S-layer gene is expressed in the host cell which codes for an unmodified S-layer protein.

41. Process as claimed in claim 40,
w h e r e i n
the unmodified S-layer protein is capable of forming
an S-layer structure that is compatible with the
recombinant S-layer protein.
42. Process as claimed in one of the claims 37-39,
w h e r e i n
no further S-layer gene is expressed in the host cell
which codes for an unmodified S-layer protein which
is capable of forming an S-layer structure that is
compatible with a recombinant S-layer protein.
43. Process as claimed in one of the claims 37-42,
w h e r e i n
a prokaryotic host cell is used.
44. Process as claimed in claim 43,
w h e r e i n
a gram-positive host cell is used.
45. Process as claimed in claim 44,
w h e r e i n
B.stearothermophilus is used.

The invention concerns a process for the recombinant production of S-layer proteins in gram-negative host cells. Furthermore the nucleotide sequence of a new S-layer gene and processes for the production of modified S-layer proteins are disclosed.

Fig.1

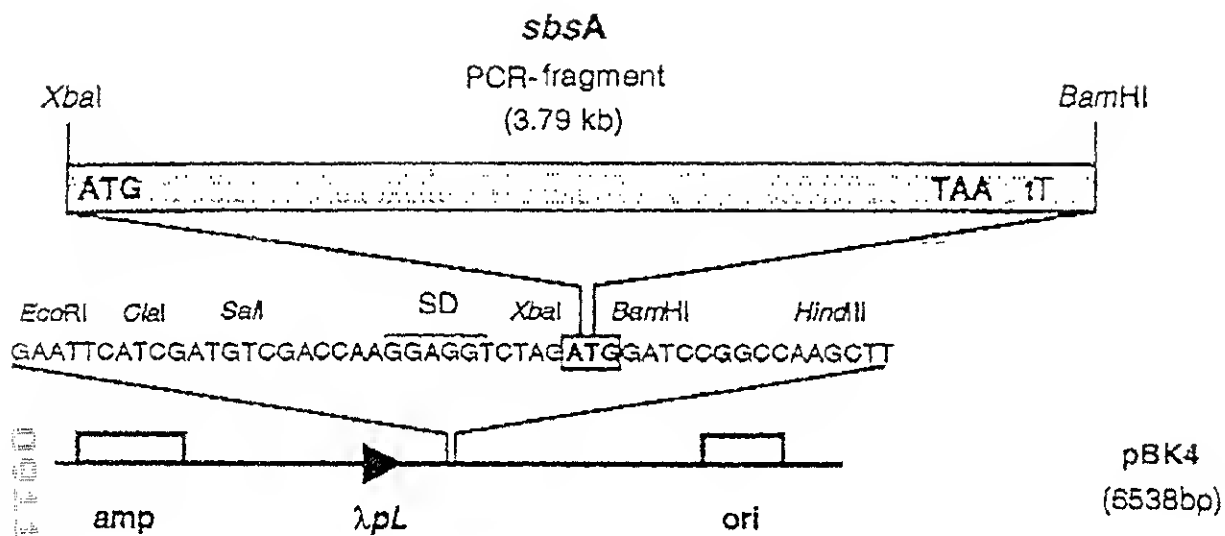
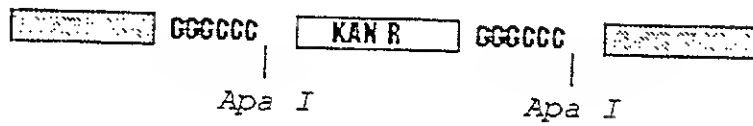


Fig. 2

A)



B)



c)

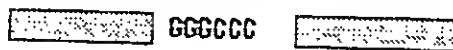
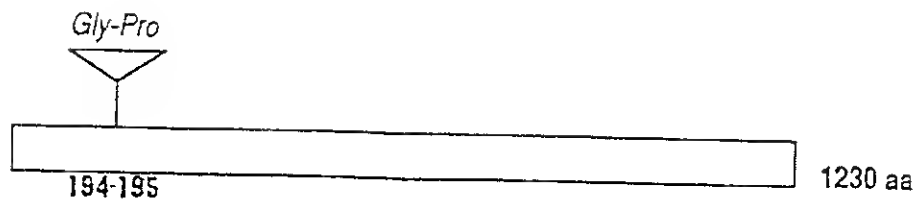


Fig.3

A)



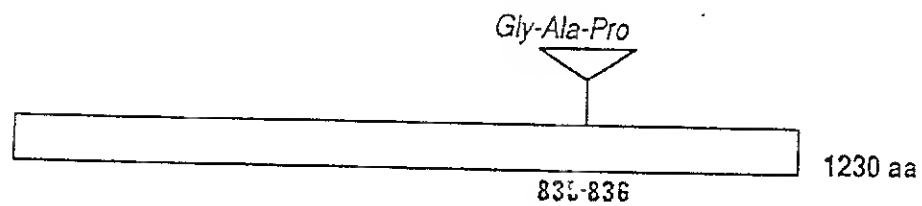
B)



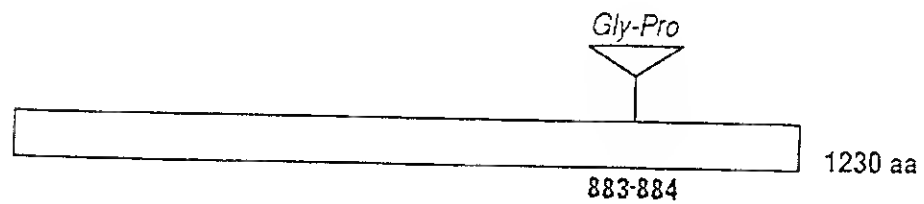
C)



D)



E)



Declaration For U.S. Patent Application

As a below named inventor, I hereby declare that:

My residence, post office address and citizenship are as stated below my name.

I believe I am the original, first and sole inventor (if only one name is listed below) or an original, first and joint inventor (if plural names are listed below) of the subject matter which is claimed and for which a patent is sought on the invention entitled
 (Insert Title) Recombinant expression of S-layer proteins

the specification of which

- (Check one of blocks 1, 2 or 3. See note A on back of this page)
1. ☐ is attached hereto.
 2. ☐ was filed on _____ as
 International PCT Application Serial No. _____
 and was amended on _____
 (if applicable)
 3. ☒ was filed on 31 July 1998 as
 U.S. Application Serial No. 09/117,447
 and was amended on _____
 (if applicable)

I hereby state that I have reviewed and understand the contents of the above-identified specification, including the claim(s), as amended by any amendment referred to above.

I acknowledge the duty to disclose information which is material to the examination of this application in accordance with Title 37, Code of Federal Regulations, §1.56(a).

I hereby claim foreign priority benefits under Title 35, United States Code, §119 of any foreign application(s) for patent or inventor's certificate listed below and have also identified below any foreign application for patent or inventor's certificate having a filing date before that of the application for which priority is claimed:

(List prior foreign applications. See note B on back of this page)	196 03 649.6	Germany	1/Feb./1996	Priority Claimed <input checked="" type="checkbox"/> Yes <input type="checkbox"/> No
(Number)	(Country)	(Day/Month/Year Filed)		
(Number)	(Country)	(Day/Month/Year Filed)		<input type="checkbox"/> Yes <input type="checkbox"/> No
(Number)	(Country)	(Day/Month/Year Filed)		<input type="checkbox"/> Yes <input type="checkbox"/> No
(Number)	(Country)	(Day/Month/Year Filed)		<input type="checkbox"/> Yes <input type="checkbox"/> No

(See Note C on back of this page)

☐ See attached list for additional prior foreign applications

I hereby claim the benefit under Title 35, United States Code, §120 of any United States application(s) or PCT International application(s) designating the United States of America listed below and, insofar as the subject matter of each of the claims of this application is not disclosed in the prior application(s) in the manner provided by the first paragraph of Title 35, United States Code, §112, I acknowledge the duty to disclose material information as defined in Title 37, Code of Federal Regulations, §1.56(a) which occurred between the filing date of the prior application and the national or PCT International filing date of this application:

(List prior U.S. Applications or PCT International applications designating the U.S.)	(Application Serial No.)	(Filing Date)	(Status) (patented, pending, abandoned)
	(Application Serial No.)	(Filing Date)	(Status) (patented, pending, abandoned)

And I hereby appoint as principal attorneys David T. Nikaido, Reg. No. 22,663; Charles M. Marmelstein, Reg. No. 25,895; George E. Oram, Jr., Reg. No. 27,931; Robert B. Murray, Reg. No. 22,980; Martin S. Postman, Reg. No. 18,570; E. Marcie Emas, Reg. No. 32,131; Michael G. Gilman, Reg. No. 19,114; Douglas H. Goldhush, Reg. No. 33,125; Kevin C. Brown, Reg. No. 32,402; Monica Chin Kitz, Reg. No. 36,105; Sharon N. Klesner, Reg. No. 36,335, and John R. Fuisz, Reg. No. 37,327.

Please direct all communications to the following address: **NIKAIDO, MARMELSTEIN, MURRAY & ORAM**
 Metropolitan Square
 655 Fifteenth Street, N.W., Suite 330 - G Street Lobby
 Washington, D.C. 20005-5701
 (202) 638-5000 Fax: (202) 638-4810

I hereby declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true; and further that these statements were made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code and that such willful false statements may jeopardize the validity of the application or any patent issued thereon.

(See Note D on back of this page)

Full name of sole or first inventor Prof. Dr. Werner Lubitz
 Inventor's signature [Signature] 30.11.98
 Residence A-1080 Wien/Vienna, Austria ATX Date
 Citizenship Austria
 Post Office Address Schönbornngasse 12/7, A-1080 Wien/Vienna, Austria

2-00 Full name of second joint inventor, if any Prof. Dr. Uwe Slevtr
Inventor's signature [Signature] 30.11.98
Residence A-1170 Wien/Vienna, Austria ATX Date
Citizenship Austria
Post Office Address Parhamerplatz 10, A-1170 Wien/Vienna, Austria

3-00 Full name of third joint inventor, if any Beatrix Kuen
Inventor's signature [Signature] 30.11.98
Residence A-1070 Wien/Vienna, Austria ATX Date
Citizenship Austria
Post Office Address Kaiserstraße 107/21, A-1070 Wien/Vienna, Austria

4-00 Full name of fourth joint inventor, if any Michaela Truppe
Inventor's signature [Signature] 30.11.98
Residence A-4222 Luftenberg, Austria ATX Date
Citizenship Austria
Post Office Address Tulpenstraße 6, A-4222 Luftenberg, Austria

5-00 Full name of fifth joint inventor, if any Stefan Howorka
Inventor's signature [Signature] 30.11.98
Residence A-1130 Wien/Vienna, Austria ATX Date
Citizenship Austria
Post Office Address Hietzinger Hauptstraße 42d/b, A-1130 Wien/Vienna, Austria

6-00 Full name of sixth joint inventor, if any Stepanka Resch
Inventor's signature [Signature] 30.11.98
Residence A-1150 Wien/Vienna, Austria ATX Date
Citizenship Austria
Post Office Address Kranzgasse 7/28, A-1150 Wien/Vienna, Austria

7-00 Full name of seventh joint inventor, if any Gerhard Schroll
Inventor's signature [Signature] 30.11.98
Residence A-1070 Wien/Vienna, Austria ATX Date
Citizenship Austria
Post Office Address Maria Hilferstraße 110/6, A-1070 Wien/Vienna, Austria

8-00 Full name of eighth joint inventor, if any Margit Sara
Inventor's signature [Signature] 30.11.98
Residence A- 2230 Gänserndorf, Austria ATX Date
Citizenship Austria
Post Office Address Watzekgasse 84, A-2230 Gänserndorf, Austria